1994

An investigation into the likelihood-based procedures for the construction of confidence intervals for the common odds ratio in K 2 x 2 contingency tables.

Sundareswary Thedchanamoorthy
University of Windsor

AN INVESTIGATION INTO THE LIKELIHOOD BASED PROCEDURES FOR THE CONSTRUCTION OF CONFIDENCE INTERVALS FOR THE COMMON ODDS RATIO IN K 2X2 CONTINGENCY TABLES

BY

SUNDARESWARY THEDCHANAMOORTHY

A Thesis

Submitted to the Faculty of Graduate Studies and Research

through the Department of Mathematics and Statistics

in Partial Fulfillment of the Requirements for

the Degree of Master of Science at the

University of Windsor

1994

ISBN 0-612-01492-4

Name S. THEECHANAMOORTHY

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

STATISTICS — SUBJECT TERM

0463 — SUBJECT CODE

U·M·I

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS
Architecture 0729
Art History 0377
Cinema 0900
Dance 0378
Fine Arts 0357
Information Science 0723
Journalism 0391
Library Science 0399
Mass Communications 0708
Music 0413
Speech Communication 0459
Theater 0465

EDUCATION
General 0515
Administration 0514
Adult and Continuing 0516
Agricultural 0517
Art 0273
Bilingual and Multicultural 0282
Business 0688
Community College 0275
Curriculum and Instruction 0727
Early Childhood 0518
Elementary 0524
Finance 0277
Guidance and Counseling 0519
Health 0680
Higher 0745
History of 0520
Home Economics 0278
Industrial 0521
Language and Literature 0279
Mathematics 0280
Music 0522
Philosophy of 0998
Physical 0523
Psychology 0525
Reading 0535
Religious 0527
Sciences 0714
Secondary 0533
Social Sciences 0534
Sociology of 0340
Special 0529
Teacher Training 0530
Technology 0710
Tests and Measurements 0288
Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS
Language
- General 0679
- Ancient 0289
- Linguistics 0290
- Modern 0291

Literature
- General 0401
- Classical 0294
- Comparative 0295
- Medieval 0297
- Modern 0298
- African 0316
- American 0591
- Asian 0305
- Canadian (English) 0352
- Canadian (French) 0355
- English 0593
- Germanic 0311
- Latin American 0312
- Middle Eastern 0315
- Romance 0313
- Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY
Philosophy 0422
Religion
- General 0318
- Biblical Studies 0321
- Clergy 0319
- History of 0320
- Philosophy of 0322

Theology 0469

SOCIAL SCIENCES
American Studies 0323
Anthropology
- Archaeology 0324
- Cultural 0326
- Physical 0327

Business Administration
- General 0310
- Accounting 0272
- Banking 0770
- Management 0454
- Marketing 0338

Canadian Studies 0385
Economics
- General 0501
- Agricultural 0503
- Commerce-Business 0505
- Finance 0508
- History 0509
- Labor 0510
- Theory 0511

Folklore 0358
Geography 0366
Gerontology 0351
History
- General 0578
- Ancient 0579
- Medieval 0581
- Modern 0582
- Black 0328
- African 0331
- Asia, Australia and Oceania 0332
- Canadian 0334
- European 0335
- Latin American 0336
- Middle Eastern 0333
- United States 0337

History of Science 0585
Law 0398
Political Science
- General 0615
- International Law and Relations 0616
- Public Administration 0617

Recreation 0814
Social Work 0452
Sociology
- General 0626
- Criminology and Penology 0627
- Demography 0938
- Ethnic and Racial Studies 0631
- Individual and Family Studies 0628
- Industrial and Labor Relations 0629
- Public and Social Welfare 0630
- Social Structure and Development 0700
- Theory and Methods 0344

Transportation 0709
Urban and Regional Planning 0999
Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES
Agriculture
- General 0473
- Agronomy 0285
- Animal Culture and Nutrition 0475
- Animal Pathology 0476
- Food Science and Technology 0359
- Forestry and Wildlife 0478
- Plant Culture 0479
- Plant Pathology 0480
- Plant Physiology 0817
- Range Management 0777
- Wood Technology 0746

Biology
- General 0306
- Anatomy 0287
- Biostatistics 0308
- Botany 0309
- Cell 0379
- Ecology 0329
- Entomology 0353
- Genetics 0369
- Limnology 0793
- Microbiology 0410
- Molecular 0307
- Neuroscience 0317
- Oceanography 0416
- Physiology 0433
- Radiation 0821
- Veterinary Science 0778
- Zoology 0472

Biophysics
- General 0786
- Medical 0760

EARTH SCIENCES
Biogeochemistry 0425
Geochemistry 0996
Geodesy 0370
Geology 0372
Geophysics 0373
Hydrology 0388
Mineralogy 0411
Paleobotany 0345
Paleoecology 0426
Paleontology 0418
Paleozoology 0985
Palynology 0427
Physical Geography 0368
Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES
Environmental Sciences 0768
Health Sciences
- General 0566
- Audiology 0300
- Chemotherapy 0992
- Dentistry 0567
- Education 0350
- Hospital Management 0769
- Human Development 0758
- Immunology 0982
- Medicine and Surgery 0564
- Mental Health 0347
- Nursing 0569
- Nutrition 0570
- Obstetrics and Gynecology 0380
- Occupational Health and Therapy 0354
- Ophthalmology 0381
- Pathology 0571
- Pharmacology 0419
- Pharmacy 0572
- Physical Therapy 0382
- Public Health 0573
- Radiology 0574
- Recreation 0575

Speech Pathology 0460
Toxicology 0383
Home Economics 0386

PHYSICAL SCIENCES

Pure Sciences

Chemistry
- General 0485
- Agricultural 0749
- Analytical 0486
- Biochemistry 0487
- Inorganic 0488
- Nuclear 0738
- Organic 0490
- Pharmaceutical 0491
- Physical 0494
- Polymer 0495
- Radiation 0754

Mathematics 0405
Physics
- General 0605
- Acoustics 0986
- Astronomy and Astrophysics 0606
- Atmospheric Science 0608
- Atomic 0748
- Electronics and Electricity 0607
- Elementary Particles and High Energy 0798
- Fluid and Plasma 0759
- Molecular 0609
- Nuclear 0610
- Optics 0752
- Radiation 0756
- Solid State 0611

Statistics 0463

Applied Sciences

Applied Mechanics 0346
Computer Science 0984

Engineering
- General 0537
- Aerospace 0538
- Agricultural 0539
- Automotive 0540
- Biomedical 0541
- Chemical 0542
- Civil 0543
- Electronics and Electrical 0544
- Heat and Thermodynamics 0348
- Hydraulic 0545
- Industrial 0546
- Marine 0547
- Materials Science 0794
- Mechanical 0548
- Metallurgy 0743
- Mining 0551
- Nuclear 0552
- Packaging 0549
- Petroleum 0765
- Sanitary and Municipal 0554
- System Science 0790

Geotechnology 0428
Operations Research 0796
Plastics Technology 0795
Textile Technology 0994

PSYCHOLOGY
General 0621
Behavioral 0384
Clinical 0622
Developmental 0620
Experimental 0623
Industrial 0624
Personality 0625
Physiological 0989
Psychobiology 0349
Psychometrics 0632
Social 0451

Approved by:

Dr. S.R. Paul, Supervisor
Department of Mathematics
and Statistics

Dr. K.Y. Fung
Department of Mathematics
and Statistics

Dr. J.M. Morrissey
Department of Computer Science

Dedicated to my Family

ABSTRACT

This study was undertaken to construct confidence intervals of the common odds ratio using several likelihood based procedures. The likelihood based procedures for the construction of confidence intervals of common odds ratio in K 2X2 contingency tables are derived. Simulations are performed to study the properties of these procedures in terms of the tail and coverage probabilities and average lengths of the confidence intervals and the results are presented. Based on the simulation results obtained in this study, it is concluded that the Bartlett method (B) is most suitable for constructing confidence interval for the common odds ratio in large sample design.

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to my supervisor Prof. S. R. Paul for his guidence and encouragement during the course of this study. I would also like to thank Professors K. Y. Fung and J. Morrissey for having served on my advisory committee.

My thanks are also due to Department of Mathematics and Statistics and the University of Windsor, for giving me finiancial support in terms of teaching assistantships and scholarships. The finiancial support from Dr. S. R. Paul through his NSERC research grant is also appreciated.

The friendship of many graduate students, faculty and the staff at the Department of Mathematics and Statistics made my carrier enjoyable and is deeply appreciated.

Finally the author expresses a very special gratitude to our family: my parents, my husband, our daughter Thulasi and son Arun. Their understanding, love and encouragement served as source of inspiration.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

The comparision of two proportions in statistics has been actively studied by researchers for many years. One approach used for comparision of two proportions is inference regarding the corresponding odds ratio, a commonly used measure of association. The inference for the odds ratio is widely used in biostatistics, such as case-control and follow-up (restrospective and prospective) studies in cancer epidemiology. In a case-control study, odds ratio is the ratio of odds of disease occurrence among the exposd group and the corresponding odds for the unexposed group. In a follow-up study, odds ratio is the ratio of odds of exposure for the disease group and the corresponding odds for the non-disease group.

The key parameter for the case-control study or for the follow-up study is the odds ratio (ψ), because it takes the same value whether it is calculated from the exposure or from the disease probabilities. In the above situation, we deal with only a 2×2 table. However, nuisance or confounding factors are involved in many studies. Confounding is defined as the distortion of a disease/exposure brought about by the association of other factors with both disease and exposure. For example, age is a confounding factor in the case of alcohol consumption and cancer. One of the most important methods known for a long time used to control the confounding factor, is to divide the sample into series of strata which are internally homogeneous with respect to the confounding factors. In such situations, the summary measure will be the common odds ratio. A full analysis of such series of 2×2 tables would be: (1) to test the homogeneity of the odds ratios in all tables; (2) once such a hypothesis is not rejected, to test the common odds ratio $\psi = 1$, that is, to test that there is no interaction between the exposure and disease and (3) if such a test fails (that is, when $\psi = 1$ is not acceptable) then to obtain the confidence interval for the common odds ratio.

Considerable amount of work has been done in this area. For example, Mantel, Brown and Byar (1977), Tarone (1985) and Paul and Donner (1989,1992) studied procedures for testing homogenity of odds ratio when the number of strata is fixed and sample size in each stratum can take any value up to infinity. Liang and Self (1985) studied procedure for testing homogenity of odds ratios in a large number of tables with sparse data in each table. Procedures for testing $\psi = 1$ were developed by Cochran (1954), Mantel and Haenzel (1959) and Mantel and Fleiss (1980).

Several point estimators for the common odds ratio exist in the literature. Woolf (1955) proposed the emprical logit estimator that behaves well for the large data but not for the sparse data. Gart (1962, 1971) developed unconditional and conditional maximum likelihood estmators. Mantel and Haenzel (1954) developed the Mantel-Haenzel (M-H) estimator. Breslow and Liang (1982) recommended a modification of the M-H estimator based on the jacknife principle. A number of simulation studies have been conducted to compare the properties (bias and precision) of various estimators of the common odds ratio (McKinlay, 1975, 1978; Lubin, 1981; Hauck, Andersen and Leahy, 1982, 1984; Jewell, 1984)

Relatively less attention has been given to confidence interval procedures for the common odds ratio. Gart (1970) gave an exact and an approximate method to construct the confidence interval for the common odds ratio. Brown (1981) studied the validity of three approximate methods developed by Cornfield (1956), Miettinen and Woolf (1955) for constructing confidence interval for the common odds ratio in a single 2×2 table. Hauck and Wallemark (1983) studied seven methods to construct the confidence interval in multiple tables. From the above study, the authors have concluded that the method using M-H estimator with Breslow's variance estimator provides coverage close to nominal. Robins, Breslow and Greendland (1986) have compared six procedures based on various estimators of the variance of the M-H estimator to construct confidence interval for the common odds ratio. Sato (1990) developed a new confidence interval procedure using the M-H estimator and its asymptotic variance.

Several likelihood based procedures for constructing the confidence interval for a parameter in the presence of nuisance parameters are available in the literature (Bartlett, 1953; Levin and Kong, 1980; Diciccio, 1990; Fraser, 1991). However, these procedures have not been used to construct the confidence interval for the common odds ratio. In this thesis, we apply several likelihood based procedures to construct confidence interval for the common odds ratio. Properties of these confidence intervals, in terms of coverage, are investigated by simulation.

In chapter 2, we review five likelihood based procedures to construct confidence interval for a parameter of interest. In chapter 3, we review maximum likelihood estimation of the common odds ratio. In chapter 4, we derive the likelihood based procedures to construct confidence interval for the common odds ratio. In chapter 5, we conduct a simulation study to investigate the properties of the various interval estimation procedures derived in chapter 4.

CHAPTER 2

A REVIEW OF LIKELIHOOD BASED PROCEDURES FOR THE CONSTRUCTION OF CONFIDENCE INTERVAL

Let $f(X;\gamma,\rho)$ be a density of a random variable X indexed by γ and ρ, where γ is the parameter of interest and $\rho = (\rho_1, \cdots, \rho_K)'$ is a vector of K nuisance parameters. Given the sample $X_1, \cdots, X_n$ denote the log-likelihood by $l(\gamma, \rho)$. Now, define the likelihood scores $\frac{\partial l}{\partial \gamma}$ and $\frac{\partial l}{\partial \rho}$. Then the maximum likelihood estimates (MLEs) of the parameters γ and $\rho = (\rho_1, \cdots, \rho_K)'$ are obtained by solving

$$\frac{\partial l}{\partial \gamma} = 0$$

and

$$\frac{\partial l}{\partial \rho_k} = 0, \qquad k = 1, \cdots, K$$

simultaneously.

2.1 Procedure based on the asymptotic properties of MLE

Denote the MLEs of the given parameters γ and $\rho = (\rho_1, \cdots, \rho_K)'$ by $\hat{\gamma}$ and $\hat{\rho}$ $= (\hat{\rho}_1, \cdots, \hat{\rho}_K)'$ respectively. The asymptotic $100(1-\alpha)\%$ confidence interval for γ is given by

$$\hat{\gamma} - \zeta\sqrt{var\hat{\gamma}} < \gamma < \hat{\gamma} + \zeta\sqrt{var\hat{\gamma}}$$

where ζ is an appropriate quantile of a standard normal random variable. The quantity $var(\hat{\gamma})$ is obtained by inverting the Fisher information matrix of $(\hat{\gamma}, \hat{\rho})$. The elements of the Fisher information matrix are the negative of the expected values of the second order mixed partial derivatives of the log-likelihood function with respect to the parameters γ and ρ. Thus, $l(X, \gamma, \rho)$ is the log-likelihood function. Then the asymptotic variance-covariance of $(\hat{\gamma}, \hat{\rho})$ is given by

$$I^{-1} = \begin{pmatrix} I_{\gamma\gamma} & I_{\gamma\rho} \\ I_{\rho\gamma} & I_{\rho\rho} \end{pmatrix}^{-1},$$

where

$$I_{\gamma\gamma} = -E(\frac{\partial^2 l}{\partial \gamma^2}),$$

$$I_{\gamma\rho} = -E(\frac{\partial^2 l}{\partial \gamma \partial \rho}),$$

$$I_{\rho\gamma} = -E(\frac{\partial^2 l}{\partial \rho \partial \gamma}),$$

and

$$I_{\rho\rho} = -E(\frac{\partial^2 l}{\partial \rho \partial \rho'}).$$

The unknown parameters in $var(\hat{\gamma})$ are then replaced by their corresponding maximum likelihood estimators. Note that $I_{\gamma\gamma}$ is a scalar, $I_{\gamma\rho}$ is a 1×K matrix, $I_{\rho\gamma}$ is a K×1 matrix and $I_{\rho\rho}$ is a K×K matrix.

2.2 Procedure based on Likelihood Ratio

Denote the unconstrained maximum log likelihood by $l(\hat{\gamma}, \hat{\rho})$ and the constrained maximum likelihood by $l(\gamma, \tilde{\rho})$, where $\tilde{\rho} = (\tilde{\rho}_1, \cdots, \tilde{\rho}_K)'$ which maximize the log-likelihood function $l(\gamma, \rho)$ for given value of γ. Then the likelihood ratio is given by

$$LR = 2\,(l(\hat{\gamma}, \hat{\rho}) - l(\gamma, \tilde{\rho}))$$

has a distribution which is approximately chi-square with one degree of freedom. Thus, the γ values that satisfy

$$LR = 2\,(l(\hat{\gamma}, \hat{\rho}) - l(\gamma, \tilde{\rho})) \leq \chi^2_{(1-\alpha)}(1)$$

are the approximate $100(1-\alpha)\%$ confidence limits for γ, where $\chi^2_{(1-\alpha)}(1)$ is the $(1-\alpha)$th quantile of a chi-squared distribution with one degree of freedom.

2.3 Procedure based on adjusted likelihood ratio

Diciccio (1988) and Diciccio, Fraser and Field (1990) developed a confidence interval procedure for the parameters of a location-scale family of distributions, where the location may be a function of several regression variables $X_1, \cdots, X_K$. Thus, if $k = 1$ we deal with the confidence interval procedure for the parameters of a two parameter distribution. Let $\rho = (\rho_1, \cdots, \rho_K)'$ be the regression parameters and γ be the scalar parameter. In many situations, inference for a scalar parameter in the presence of nuisance parameters requires pivotal quantities. From Diciccio (1988) the pivotal statistics are,

$$P_k = \frac{\rho_k - \hat{\rho}_k}{\hat{\gamma}}, \qquad k = 1, \cdots, K$$

and

$$P_{K+1} = log(\frac{\gamma}{\hat{\gamma}}).$$

where $\hat{\rho}$ and $\hat{\gamma}$ are the MLE's of $\rho = (\rho_1, \cdots, \rho_K)'$ and γ. Therefore, the log-likelihood $l(\gamma, \rho)$ can be written in terms of a vector of pivotals $P = (P_1, \cdots, P_{K+1})'$. We denote this as $l(P)$. It is obvious that the likelihood $l(P)$ attains its maximum value $l(0)$ at $P_k = 0, \quad k = 1, \cdots, K+1$. Suppose the kth parameter is of interest, then the associated pivotal is P_k and the corresponding likelihood ratio (LR_k) is

$$LR_k = 2\left[l(0) - l(\tilde{P}(P_k))\right]$$

where $l(\tilde{P}(P_k))$ is the maximized log-likelihood function for a given value of P_k. The statistic LR_k is approximately distributed as chi-square with one degree of freedom. Now define the signed root of the likelihood ratio by

$$SR_k = -\sqrt{LR_k}, \qquad P_k < 0$$

and

$$SR_k = +\sqrt{LR_k}, \qquad P_k > 0.$$

The distribution of SR_k can be approximated by the standard normal distribution, which has an error of order $n^{-\frac{1}{2}}$. That is

$$Pr(P_k \leq p_k) = \Phi(SR_k) + 0(n^{\frac{-1}{2}}),$$

where Φ is the distribution function of a standard normal random variable. Many researchers including Brandorff-Nielsen (1986), Diciccio (1984, 1988), Efron (1985) and McCullagh (1984; 1987) studied on further reduction of error and concluded that mean and variance adjustment to the approximate standard normal distribution of the signed root likelihood ratio statistics reduces the error to the order $n^{-\frac{3}{2}}$. Thus,

$$Pr(P_k \leq p_k) = \Phi(\frac{SR_k - \mu_k}{\sigma_k}) + 0(n^{-\frac{3}{2}})$$

where μ_k and σ_k^2 are the mean and variance of SR_k, respectively. The above equation can not be used as the exact values of mean and variance are not available. However, in principle, they can be sufficiently well approximated such that the above equation remains valid. Diciccio, Field and Fraser (1990) presented a procedure whereby the mean and variance adjustments in the above equation can be achieved using a simple formula that involves only first and second order partial derivatives.

The general form of the approximation is

$$Pr(P_k \leq p_k) = \Phi(SR_k) + \phi(SR_k)\left[\frac{1}{SR_k} + \frac{|I|^{\frac{1}{2}}}{l_1(\tilde{P}(P_k)) \times |I^*|^{\frac{1}{2}}}\right] + 0(n^{\frac{-3}{2}})$$

where I is the observed information matrix of order $(K+1) \times (K+1)$ with P_k, $k = 1, ..., K, K+1$ being replaced by zero. I^* is the submatrix of I corresponding to $(P_1, \cdots, P_{k-1}, P_{k+1}, \cdots, P_{K+1})$ with P_j , $j = 1, ...K, K+1$, $j \neq k$ being replaced by its maximum likelihood estimate for given value of P_k. $\mid I \mid^{\frac{1}{2}}$ and $\mid I^* \mid^{\frac{1}{2}}$ are the square roots of the determinants of the matrices I and I^* respectively for $k = 1, \cdots, K+1$ and $l_1(\tilde{P}(P_k) = \frac{\partial l}{\partial P_k}\Big|_{P_k=P_k}$, $P_j = \tilde{P}_j$, $j = 1, ...K, K+1, j \neq k$.

When $k = 1$, the above approximation reduces to

$$Pr(P \leq p) = \Phi(SR) + \phi(SR)\left[\frac{1}{SR} + \frac{(-l_2(0))^{\frac{1}{2}}}{l_1(p)}\right] + 0(n^{\frac{-3}{2}})$$

where

$$l_1(x) = \frac{\partial^{(i)} l(x)}{\partial x^i}$$

and ϕ is the density function of N(0,1). Thus, $100(1-\alpha)\%$ approximate lower and upper confidence limits for the given pivotal (kth) are obtained by solving

$$Pr(P_k \leq p_k) = \frac{\alpha}{2}$$

and

$$Pr(P_k \leq p_k) = 1 - \frac{\alpha}{2}.$$

Hence, the confidence limits for the kth parameter of interest can be obtained from the pivotal limits.

2.4 Bartlett's procedure based on the likelihood score

Bartlett (1953) showed that, in the case of "nuisance parameter" there is an alternative to maximum likelihood estimator $\hat{\gamma}$ and is given by

$$T(\gamma) = \left(\frac{\partial l}{\partial \gamma} - I_{\gamma\rho} I_{\rho\rho}^{-1} \frac{\partial l}{\partial \rho}\right)$$

with variance

$$I_{\gamma\gamma\cdot\rho} = I_{\gamma\gamma} - I_{\gamma\rho}.I_{\rho\rho}^{-1}.I_{\rho\gamma}.$$

Also, he showed that $\frac{T(\gamma)}{\sqrt{I_{\gamma\gamma\cdot\rho}}}$ is a standardized normal variable. That is,

$$\frac{T(\gamma)}{\sqrt{I_{\gamma\gamma\cdot\rho}}} \sim N(0,1).$$

An approximate $100(1-\alpha)\%$ confidence interval for γ can then be obtained by solving

$$\frac{T(\gamma)}{\sqrt{I_{\gamma\gamma\cdot\rho}}} = \pm Z_{\frac{\alpha}{2}}$$

where $Z_{\frac{\alpha}{2}}$ is the appropriate quantile of the standard normal random variate.

2.5 Bartlett's procedure corrected for bias and skewness

When the nuisance parameters ρ in $T(\gamma)$ are replaced by their corresponding maximum likelihod estimates, the statistic $T(\gamma)$ involves a bias of $O(n^{-\frac{1}{2}})$ and is given by

$$Bias = B(T(\gamma)) = -\frac{1}{2}trace\left(I_{\rho\rho}^{-1}\left(E(\frac{\partial^3 l}{\partial\gamma\partial\rho\partial\rho'}) + 2\frac{\partial I_{\gamma\rho}}{\partial\rho}\right)\right) + \frac{1}{2}trace\left(I_{\rho\rho}^{-1}M\right),$$

where

$$M_j = \left(E(\frac{\partial^3 l}{\partial\rho_j\partial\rho\partial\rho'}) + 2\frac{\partial I_{\rho\rho}}{\partial\rho_j}\right) I_{\rho\rho}^{-1} I_{\rho\gamma}, \qquad j = 1,\cdots,K.$$

See, Bartlett (1953), Levin and Kong (1990).

Skewness or the third cumulant of $T(\gamma)$ to the order of $n^{-\frac{3}{2}}$ is obtained for $s = t = q = 1, ..., K$, as

$$K_3(\gamma) = 2E(\frac{\partial^3 l}{\partial\gamma^3}) + 3\frac{\partial I_{\gamma\gamma}}{\partial\gamma}$$

$$-3\sum_{s=1}^{K} f_s\left(2E(\frac{\partial^3 l}{\partial\gamma^2\partial\rho_s}) + 2E(\frac{\partial I_{\gamma\rho_s}}{\partial\gamma}) + \frac{\partial I_{\gamma\gamma}}{\partial\rho_s}\right)$$

$$+3\sum_{s}\sum_{t} f_s f_t\left(2E(\frac{\partial^3 l}{\partial\gamma\partial\rho_s\partial\rho_t}) + \frac{\partial I_{\rho_s\rho_t}}{\partial\gamma} + \frac{\partial I_{\gamma\rho_t}}{\partial\rho_s} + \frac{\partial I_{\gamma\rho_s}}{\partial\rho_t}\right)$$

$$-\sum_{s}\sum_{t}\sum_{q} f_s f_t f_q\left(2E(\frac{\partial^3 l}{\partial\rho_s\partial\rho_t\partial\rho_q}) + \frac{\partial I_{\rho_s\rho_t}}{\partial\rho_q} + \frac{\partial I_{\rho_q\rho_t}}{\partial\rho_s} + \frac{\partial I_{\rho_q\rho_s}}{\partial\rho_t}\right)$$

where $f = (f_1,\cdots,f_K)' = I_{\gamma\rho}I_{\rho\rho}^{-1}$.

The statistic $T(\gamma)$ corrected for skewness and bias are a better approximation for the normal distribution. Therefore, a more accurate $100(1-\alpha)\%$ confidence interval for γ can be obtained by solving

$$\frac{T(\gamma)}{\sqrt{I_{\gamma\gamma\cdot\rho}}} - \frac{B(T(\gamma))}{\sqrt{I_{\gamma\gamma\cdot\rho}}} - \frac{K_3(\gamma)(Z^2_{\frac{\alpha}{2}} - 1)}{6I^{\frac{3}{2}}_{\gamma\gamma\cdot\rho}} = \pm Z_{\frac{\alpha}{2}}.$$

CHAPTER 3

NOTATIONS AND ESTIMATIONS OF COMMON ODDS RATIO

3.1 Notations

Consider K pairs of mutually independent binomial variates X_{1k}, X_{2k} with corresponding parameters p_{1k}, p_{2k} and sample sizes N_{1k}, N_{2k}, where $k = 1, \cdots, K$.

$$X_{1k} \sim B(N_{1k}, p_{1k})$$

and

$$X_{2k} \sim B(N_{2k}, p_{2k}).$$

Thus the data for the kth table or for the kth pair or the kth stratum are

		1(success)	2(failure)	
group	1	X_{1k}	$N_{1k} - X_{1k}$	N_{1k}
	2	X_{2k}	$N_{2k} - X_{2k}$	N_{2k}
		T_k	$N_k - T_k$	N_k

and the corresponding table of probabilities for the kth stratum are

		1(success)	2(failure)
group	1	p_{1k}	q_{1k}
	2	p_{2k}	q_{2k}

where $p_{1k} + q_{1k} = 1$ and $p_{2k} + q_{2k} = 1$. Thus, $q_{1k} = 1 - p_{1k}$ and $q_{2k} = 1 - p_{2k}$. The odds ratio for the kth table (stratum) is

$$\psi_k = \frac{p_{1k} q_{2k}}{p_{2k} q_{1k}}. \tag{3.1}$$

The alternative name cross product ratio is used for odds ratio as it is equal to the ratio of the products $p_{1k}q_{2k}$ and $p_{2k}q_{1k}$; the probabilities from diagonally opposite cells. The odds ratio can be any nonnegative number. In other words

$$0 < \psi_k < \infty.$$

The odds ratio does not change values when the orientation of the table is reversed or when the rows become the columns and vice versa. Therefore, it is not necessary to identify the clssification as the response in order to calculate odds ratio. It is sometimes more convenient to use $log(\psi_k)$, the natural logarithm of ψ_k. Because the odds ratio is symmetric about this value, reversal of rows or columns changes only the sign. In this study, we consider only the case where the odds ratio is the same in all strata (tables). That is,

$$\psi_k = \psi, \qquad 0 < \psi < \infty$$

for all $k = 1, ..., K$.

3.2 Unconditional maximum likelihood estimator

The distributions of X_{1k} and X_{2k} are binomials with indices N_{1k} and N_{2k} and probabilities p_{1k} and p_{2k} respectively. The likelihood L, dropping the combinatorial terms, is

$$L \propto \prod_{k=1}^{K} (p_{1k})^{X_{1k}} (q_{1k})^{N_{1k}-X_{1k}} (p_{2k})^{X_{2k}} (q_{2k})^{N_{2k}-X_{2k}}.$$

Using equation 3.1 with $\psi_k = \psi$, we have

$$p_{2k} = \frac{p_{1k}}{\psi q_{1k} + p_{1k}}$$

and

$$q_{2k} = 1 - p_{2k}.$$

So the likelihood

$$L \propto \prod_{k=1}^{K} (p_{1k})^{X_{1k}} (q_{1k})^{N_{1k}-X_{1k}} (\frac{p_{1k}}{\psi q_{1k} + p_{1k}})^{X_{2k}} (\frac{\psi q_{1k}}{\psi q_{1k} + p_{1k}})^{N_{2k}-X_{2k}}$$

and the log-likelihood l is

$$l = C + \sum_{k=1}^{K} \left(X_{1k} log(\frac{p_{1k}}{q_{1k}}) + N_{1k} log q_{1k} + X_{2k} log \frac{p_{1k}}{\psi q_{1k}} + N_{2k} log \frac{\psi q_{1k}}{\psi q_{1k} + p_{1k}} \right),$$

$$= C + \sum_{k=1}^{K} \left((X_{1k} + X_{2k}) log(\frac{p_{1k}}{q_{1k}}) + N_{1k} log q_{1k} + N_{2k} log \frac{\psi}{\psi + \frac{p_{1k}}{q_{1k}}} - X_{2k} log \psi \right),$$

where, C is a constant independent of the parameters ψ, p_{1k}, q_{1k}. Now, $T_k = X_{1k} + X_{2k}$. Define $\rho_k = log\frac{p_{1k}}{q_{1k}}$, $\gamma = log\psi$. Then $\frac{p_{1k}}{q_{1k}} = e^{\rho_k}$, $p_{1k} = \frac{e^{\rho_k}}{1+e^{\rho_k}} \Rightarrow q_{1k} = \frac{1}{1+e^{\rho_k}}$. The log-likelihood l can be written as

$$l = C + \sum_{k=1}^{K} (T_k \rho_k + (N_{2k} - X_{2k})\gamma - N_{1k} log(1 + e^{\rho_k}) - N_{2k} log(e^{\gamma} + e^{\rho_k})). \tag{3.2}$$

The log-likelihood involves the $K + 1$ parameters γ and ρ_k, $k = 1, \cdots, K$. The maximum log-likelihood estimators of γ and ρ_k ($k = 1, ..., K$) are obtained by maximizing the log-likelihood (3.2) directly by using IMSL subroutine DUMINF (IMSL.LIB 1989) . Denote the MLEs of ρ_k and γ be $\hat{\rho}_k$ and $\hat{\gamma}_u$. These are the unconditional maximum likelihood estimators.

3.3 Conditional maximum likelihood

When working with the increasing strata case, the principal distribution of interest will be that of X_{1k}, given T_k, N_{1k} and N_{2k}. As originally noted by Fisher (1935), this distribution is the extended (or noncentral) hypergeometric distribution (Harkness, 1965) which is given by

$$f(X_{1k}|T_k, N_{1k}, N_{2k}) = \frac{\binom{N_{1k}}{X_{1k}} \binom{N_{2k}}{T_k - X_{1k}} \psi^{X_{1k}}}{\sum_{u=a_k}^{b_k} \binom{N_{1k}}{u} \binom{N_{2k}}{T_k - u} \psi^u}$$

where a_k =max (0, $T_k - N_{2k}$), b_k = min (T_k ,N_{1k}).

Then the joint likelihood is

$$L = \prod_{k=1}^{K} \frac{\binom{N_{1k}}{X_{1k}}\binom{N_{2k}}{T_k - X_{1k}}\psi^{X_{1k}}}{\sum_{u=a_k}^{b_k}\binom{N_{1k}}{u}\binom{N_{2k}}{T_k - u}\psi^{u}}.$$

By reparametrization of $\psi = e^{\gamma}$, we have

$$L = \prod_{k=1}^{K} \frac{\binom{N_{1k}}{X_{1k}}\binom{N_{2k}}{T_k - X_{1k}}e^{\gamma X_{1k}}}{\sum_{u=a_k}^{b_k}\binom{N_{1k}}{u}\binom{N_{2k}}{T_k - u}e^{\gamma u}}.$$

From this, the log-likelihood can be written as

$$l = \sum_{k=1}^{K} (X_{1k}\gamma + C_k - log f_k(\gamma))$$

where

$$f_k(\gamma) = \sum_{X_{1k}=a_k}^{b_k} \binom{N_{1k}}{X_{1k}}\binom{N_{2k}}{T_k - X_{1k}}e^{\gamma X_{1k}}$$

and C_k is a constant independent of the parameter γ.

The partial derivative of l with respect to the parameter γ is

$$\frac{\partial l}{\partial \gamma} = \sum_{k=1}^{K}\left(X_{1k} - \frac{1}{f_k(\gamma)}\frac{\partial f_k(\gamma)}{\partial \gamma}\right).$$

Now,

$$E(X_{1k};\gamma) = \sum_{X_{1k}=b_{1k}}^{b_k} X_{1k}\frac{\binom{N_{1k}}{X_{1k}}\binom{N_{2k}}{T_k - X_{1k}}e^{\gamma X_{1k}}}{f_k(\gamma)}.$$

Further,

$$\frac{\partial f_k(\gamma)}{\partial \gamma} = \sum_{X_{1k}=a_k}^{b_k} X_{1k}\binom{N_{1k}}{X_{1k}}\binom{N_{2k}}{T_k - X_{1k}}e^{\gamma X_{1k}}.$$

Thus,

$$E(X_{1k};\gamma) = \frac{1}{f_k(\gamma)}\frac{\partial f_k(\gamma)}{\partial \gamma}$$

and

$$\frac{\partial l}{\partial \gamma} = \sum_{k=1}^{K} X_{1k} - \sum_{k=1}^{K} E(X_{1k};\gamma).$$

But, for the maximum likelihood estimator of γ,

$$\frac{\partial l}{\partial \gamma} = 0.$$

Hence, the maximum likelihood estimator of γ is obtained by solving the equation

$$\sum_{k=1}^{K} X_{1k} = \sum_{k=1}^{K} E(X_{1k}; \gamma).$$

This equation can be solved by IMSL subroutine ZBREN. Denote the maximum likelihood estimator of γ by $\hat{\gamma}_c$. This is a conditional maximum likelihood estimator.

CHAPTER 4

LIKELIHOOD BASED CONFIDENCE INTERVAL PROCEDURE FOR THE COMMON ODDS RATIO

4.1 Confidence interval estimation based on the conditional maximum likelihood estimator of the common odds ratio.

4.1.1 Procedure based on the asymptotic properties of MLE

From the definition of variance,

$$Var(X_{1k};\gamma) = E(X_{1k}^2;\gamma) - \left(E(X_{1k};\gamma)\right)^2,$$

where

$$E(X_{1k};\gamma) = \sum_{X_{1k}=a_k}^{b_k} X_{1k} \frac{\binom{N_{1k}}{X_{1k}}\binom{N_{2k}}{T_K - X_{1k}} e^{\gamma X_{1k}}}{f_k(\gamma)}$$

and

$$E(X_{1k}^2;\gamma) = \sum_{X_{1k}=a_k}^{b_k} X_{1k}^2 \frac{\binom{N_{1k}}{X_{1k}}\binom{N_{2k}}{T_K - X_{1k}} e^{\gamma X_{1k}}}{f_k(\gamma)}.$$

From chapter 3 we have

$$\frac{\partial l}{\partial \gamma} = \sum_{k=1}^{K} X_{1k} - \sum_{k=1}^{K} \frac{1}{f_k(\gamma)} \frac{\partial f_k(\gamma)}{\partial \gamma} \tag{4.1}$$

so the second derivative with respect to γ is

$$\frac{\partial^2 l}{\partial \gamma^2} = -\left(\sum_{k=1}^{K} \frac{1}{f_k(\gamma)} \frac{\partial^2 f_k(\gamma)}{\partial \gamma^2} - \sum_{k=1}^{K} \frac{1}{f_k^2(\gamma)} (\frac{\partial f_k(\gamma)}{\partial \gamma})^2 \right).$$

From $f_k(\gamma)$, we have

$$\frac{\partial f_k(\gamma)}{\partial \gamma} = \sum_{X_{1k}=a_k}^{b_k} X_{1k} \binom{N_{1k}}{X_{1k}} \binom{N_{2k}}{T_k - X_{1k}} e^{\gamma X_{1k}}$$

and

$$\frac{\partial^2 f_k(\gamma)}{\partial \gamma^2} = \sum_{X_{1k}=a_k}^{b_k} X_{1k}^2 \binom{N_{1k}}{X_{1k}} \binom{N_{2k}}{T_k - X_{1k}} e^{\gamma X_{1k}}.$$

So

$$E(X_{1k};\gamma) = \frac{1}{f_k(\gamma)} \frac{\partial f_k(\gamma)}{\partial \gamma}$$

and

$$E(X_{1k}^2;\gamma) = \frac{1}{f_k(\gamma)} \frac{\partial^2 f_k(\gamma)}{\partial \gamma^2}.$$

Thus,

$$\frac{\partial^2 l}{\partial \gamma^2} = -\left(\sum_{k=1}^{K} E(X_{1k}^2;\gamma) - \sum_{k=1}^{K} (E(X_{1k};\gamma))^2 \right)$$

$$= -\sum_{k=1}^{K} Var(X_{1k};\gamma).$$

Now, the asymptotic variance of the conditional maximum likelihood estimator $\hat{\gamma}_c$ is

$$Var(\hat{\gamma}_c) = -\frac{1}{E(\frac{\partial^2 l}{\partial \gamma^2})}.$$

Therefore,

$$Var(\hat{\gamma}_c) = \frac{1}{\sum_{k=1}^{K} Var(X_{1k};\gamma)}.$$

Thus, an approximate confidence interval estimation for $\hat{\gamma}_c$ is obtained as

$$\hat{\gamma}_c \pm Z_{\frac{\alpha}{2}} \sqrt{Var\hat{\gamma}_c}$$

which can be written as

$$\hat{\gamma}_c \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{1}{\sum_{k=1}^{k} var(X_{1k};\gamma)}} \qquad (4.2)$$

Denote the estimates of the lower and upper limits of the confidence interval for γ obtained from (4.2) by $\hat{\gamma}_{McL}$ and $\hat{\gamma}_{McU}$. Then an approximate confidence interval for ψ, by using the conditional maximum likelihood estimator of γ, is

$$e^{\hat{\gamma}_{McL}}$$

and

$$e^{\hat{\gamma}_{McU}}.$$

Denote these by $\hat{\psi}_{McL}$ and $\hat{\psi}_{McU}$.

4.1.2 Procedure based on likelihood ratio

From chapter 3, the conditional log-likeihood is

$$l(\gamma) = \sum_{k=1}^{K} (X_{1k}\gamma + C_k - log f_k(\gamma))$$

and the maximized log-likelihood, using the maximum likelihood estimator $\hat{\gamma}_c$ is

$$l(\hat{\gamma}_c) = \sum_{k=1}^{K} (X_{1k}\hat{\gamma}_c + C_k - log f_k(\hat{\gamma}_c)).$$

Thus,

$$\begin{aligned} LR &= 2\,[l(\hat{\gamma}_c) - l(\gamma)] \\ &= 2\sum_{k=1}^{K} X_{1k}(\hat{\gamma}_c - \gamma) + log\frac{f_k(\gamma)}{f_k(\hat{\gamma}_c)}. \end{aligned}$$

The confidence limit for γ using the likelihood ratio procedure is obtained by solving

$$2\sum_{k=1}^{K}\left[X_{1k}(\hat{\gamma}_c - \gamma) + log\frac{f_k(\gamma)}{f_k(\hat{\gamma}_c)}\right] = \chi^2_{1-\alpha}(1) \tag{4.3}$$

This equation can be solved by using IMSL subroutine ZREAL. Denote the lower and upper limits of the confidence interval for γ, obtained from (4.3) by $\hat{\gamma}_{LcL}$ and $\hat{\gamma}_{LcU}$. Then the estimators for the lower and the upper limit of the confidence interval for ψ, using the likelihood ratio procedure, are

$$\hat{\psi}_{LcL} = e^{\hat{\gamma}_{LcL}}$$

and

$$\hat{\psi}_{LcU} = e^{\hat{\gamma}_{LcU}}.$$

4.1.3 Procedure based on adjusted likelihood ratio

According to the procedure developed by Diciccio, Field and Fraser (1990), as discussed in chapter 2, odds ratio is the only parameter of interest in the conditional approach. Therefore, the pivotal statistic is

$$P = log\frac{\psi}{\hat{\psi}}.$$

By reparametrization of $log\psi = \gamma$, we have

$$P = \gamma - \hat{\gamma}.$$

Therefore,

$$\gamma = P + \hat{\gamma},$$

where $\hat{\gamma} = \hat{\gamma}_c$=conditional maximum likelihood estimator of γ. From chapter 3, the conditional log-likelihood is

$$l(\gamma) = \sum_{k=1}^{K} (X_{1k}\gamma + C_k - logf_k(\gamma)).$$

Then, the corresponding log-likelihood,using the pivotal P, is

$$l(P) = \sum_{k=1}^{K} (X_{1k}(P + \hat{\gamma}) + C_k - logf_k(P + \hat{\gamma})).$$

When $P = 0$, the corresponding log-likelihood is

$$l(0) = \sum_{k=1}^{K} (X_{1k}\hat{\gamma} + C_k - logf_k(\hat{\gamma})).$$

Hence, the likelihood ratio statistic is

$$LR = 2\,[l(0) - l(P)]$$

$$= 2\sum_{k=1}^{K} \left(X_{1k}(-P) + log\frac{f_k(\hat{\gamma} + P)}{f_k(\hat{\gamma})} \right)$$

From chapter 2, the signed root likelihood ratio is

$$SR = -\sqrt{LR}, \qquad if \quad P < 0$$

and

$$SR = \sqrt{LR}, \qquad if \quad P > 0.$$

Now, from the log-likelihood, involving the pivotal, we have

$$l_1(P) = \frac{\partial l(P)}{\partial P}$$

$$= \sum_{k=1}^{K} X_{1k} - \sum_{k=1}^{K} E(X_{1k}; \hat{\gamma} + P)$$

and

$$l_2(0) = -\sum_{k=1}^{K} Var(X_{1k}; \hat{\gamma}).$$

Follwing the procedure in section 2.3, the marginal tail probability for the pivotal P can be given as

$$Pr(P \leq p) = \Phi(SR) + \phi(SR)\left(\frac{1}{SR} + \frac{\sqrt{-l_2(0)}}{l_1(p)}\right) + O(n^{\frac{-3}{2}}).$$

Hence, the $100(1-\alpha)\%$ approximate lower and upper confidence limits are obtained by solving

$$Pr(P \leq p) = \Phi(SR) + \phi(SR)\left(\frac{1}{SR} + \frac{\sqrt{-l_2(0)}}{l_1(p)}\right) = \frac{\alpha}{2}$$

and

$$Pr(P \leq p) = \Phi(SR) + \phi(SR)\left(\frac{1}{SR} + \frac{\sqrt{-l_2(0)}}{l_1(p)}\right) = 1 - \frac{\alpha}{2}$$

respectively. Denote these as P_L and P_U. Therefore, the corresponding lower and upper limits of γ using Diciccio's procedure are $\hat{\gamma} + P_L$ and $\hat{\gamma} + P_U$ respectively. Then the lower and the upper limits of the confidence interval for ψ are

$$\hat{\psi}_{DcL} = e^{\hat{\gamma} + P_L}$$

and

$$\hat{\psi}_{DcU} = e^{\hat{\gamma}+P_U}.$$

4.1.4 Bartlett's procedure based on likelihood score

In this approach (conditional), we have determined the likelihood, using only the parameter of interest γ (or ψ). From chapter 2, the alternative to the maximum likelihood estimate $\hat{\gamma}_c$, is

$$T(\gamma) = \frac{\partial l}{\partial \gamma}$$

with variance $I_{\gamma\gamma}$. Thus, an approximate confidence interval for γ can be obtained by solving

$$\frac{T(\gamma)}{\sqrt{I_{\gamma\gamma}}} = \pm Z_{\frac{\alpha}{2}}$$

where $Z_{\frac{\alpha}{2}}$ is the appropriate quantile of the standard normal distribution. From Fisher information matrix

$$I_{\gamma\gamma} = -E(\frac{\partial^2 l}{\partial \gamma^2})$$

and also from section 4.1.1 we have

$$\frac{\partial l}{\partial \gamma} = \sum_{k=1}^{K} (X_{1k} - E(X_{1k};\gamma))$$

and

$$\frac{\partial^2 l}{\partial \gamma^2} = -\sum_{k=1}^{K} Var(X_{1k};\gamma).$$

Therefore, an approximate confidence interval for γ can be obtained by solving

$$\frac{\sum_{k=1}^{K} (X_{1k} - E(X_{1k};\gamma))}{\sqrt{\sum_{k=1}^{K} Var(X_{1k};\gamma)}} = \pm Z_{\frac{\alpha}{2}}.$$

This equation can be solved by using IMSL subroutine ZBREN. Denote the lower limit and the upper limit of γ obtained by Bartlett's procedure by $\hat{\gamma}_{BcL}$ and $\hat{\gamma}_{BcU}$. Then the corresponding lower and upper limits for the confidence interval of the odds ratio ψ are

$$\hat{\psi}_{BcL} = e^{\hat{\gamma}_{BcL}}$$

and

$$\hat{\psi}_{BcU} = e^{\hat{\gamma}_{BcU}}.$$

4.1.5 Bartlett's procedure corrected for bias and skewness

Since the conditional likelihood involves only one parameter γ, from chapter 2,

$$B\left(T(\gamma)\right) = 0.$$

The third cumulant for the alternative to the maximum lkelihood estimate for a single parameter $\hat{\gamma}_c$, is

$$K_3\left(T(\gamma)\right) = 2E(\frac{\partial^3 l}{\partial\gamma^3}) + 3\frac{\partial I_{\gamma\gamma}}{\partial\gamma}.$$

Therefore, the $100(1-\alpha)\%$ approximate confidence interval for γ is obtained by solving

$$T(\gamma) - \frac{K_3(\gamma)(Z^2_{\frac{\alpha}{2}} - 1)}{6I^{\frac{3}{2}}_{\gamma\gamma}} = \pm Z_{\frac{\alpha}{2}}.$$

Applying

$$I_{\gamma\gamma} = -E(\frac{\partial^2 l}{\partial\gamma^2}),$$

the partial derivative of $I_{\gamma\gamma}$ with respect to γ is

$$\frac{\partial I_{\gamma\gamma}}{\partial\gamma} = -E(\frac{\partial^2 l}{\partial\gamma^2}\frac{\partial l}{\partial\gamma}) - E(\frac{\partial^3 l}{\partial\gamma^3}).$$

From section 4.1.1

$$\frac{\partial^2 l}{\partial\gamma^2} = -\sum_{k=1}^{K} Var(X_{1k};\gamma).$$

Therefore,

$$-E(\frac{\partial^2 l}{\partial\gamma^2}\frac{\partial l}{\partial\gamma}) = \left(\sum_{k=1}^{K} Var(X_{1k};\gamma)\right) E(\frac{\partial l}{\partial\gamma}).$$

But under regularity conditions

$$E(\frac{\partial l}{\partial \gamma}) = 0.$$

Therefore,

$$-E(\frac{\partial^2 l}{\partial \gamma^2}\frac{\partial l}{\partial \gamma}) = 0.$$

Thus,

$$\frac{\partial I_{\gamma\gamma}}{\partial \gamma} = -E(\frac{\partial^3 l}{\partial \gamma^3}).$$

and

$$K_3\,(T(\gamma)) = -E(\frac{\partial^3 l}{\partial \gamma^3}).$$

But from section 4.1.1

$$\frac{\partial^2 l}{\partial \gamma^2} = -\left(\sum_{k=1}^{K}\frac{1}{f_k(\gamma)}\frac{\partial^2 f_k(\gamma)}{\partial \gamma^2} - \sum_{k=1}^{K}\frac{1}{f_k^2(\gamma)}(\frac{\partial f_k(\gamma)}{\partial \gamma})^2\right). \tag{4.4}$$

Hence, the third derivatives of $f_k\ (\gamma)$ with respect to γ is

$$\frac{\partial^3 l}{\partial \gamma^3} = \sum_{k=1}^{K}\left[\left(\frac{1}{f_k^2(\gamma)}(\frac{\partial f_k(\gamma)}{\partial \gamma})(\frac{\partial^2 f_k(\gamma)}{\partial \gamma^2}\right) - \left(\frac{2}{f_k^3(\gamma)}(\frac{\partial f_k(\gamma)}{\partial \gamma})(\frac{\partial f_k(\gamma)}{\partial \gamma})^2\right)\right]$$

$$-\sum_{k=1}^{K}\left[\left(\frac{1}{f_k(\gamma)}(\frac{\partial^3 f_k(\gamma)}{\partial \gamma^3})\right) - \left((2)\frac{1}{f_k(\gamma)^2}(\frac{\partial f_k(\gamma)}{\partial \gamma})(\frac{\partial^2 f_k(\gamma)}{\partial \gamma^2})\right)\right].$$

But from section 4.1.1, we have

$$E(X_{1k};\gamma) = \frac{1}{f_k(\gamma)}\frac{\partial f_k(\gamma)}{\partial \gamma}$$

and

$$E(X_{1k}^2;\gamma) = \frac{1}{f_k(\gamma)}\frac{\partial^2 f_k(\gamma)}{\partial \gamma^2}.$$

Furthermore,

$$E(X_{!k}^3;\gamma) = \frac{1}{f_k(\gamma)}\frac{\partial^3 f_k(\gamma)}{\partial \gamma^3}.$$

Hence,

$$\frac{\partial^3 l}{\partial \gamma^3} = \sum_{k=1}^{K} 3E(X_{1k};\gamma)E(X_{1k}^2;\gamma) - \sum_{k=1}^{K} 2\left(E(X_{1k};\gamma)\right)^3 - \sum_{k=1}^{K} E(X_{1k}^3;\gamma)$$

and

$$E(\frac{\partial^3 l}{\partial \gamma^3}) = \sum_{k=1}^{K} 3E(X_{1k};\gamma)E(X_{1k}^2;\gamma) - \sum_{k=1}^{K} 2\left(E(X_{1k};\gamma)\right)^3 - \sum_{k=1}^{K} E(X_{1k}^3;\gamma).$$

Therefore, the third cumulant of $T(\gamma)$ is

$$K_3(T(\gamma)) = -\sum_{k=1}^{K} 3E(X_{1k},\gamma)E(X_{1k}^2,\gamma) + \sum_{k=1}^{K} 2\left(E(X_{1k},\gamma)\right)^3 + \sum_{k=1}^{K} E(X_{1k}^3,\gamma) \qquad (4.5)$$

By using the the values for $T(\gamma)$ and $I_{\gamma\gamma}$ and $K_3(T(\gamma))$ from equations (4.1), (4.4) and (4.5) the confidence interval for γ is obtained as

$$T(\gamma) - \frac{K_3(T(\gamma))}{6I_{\gamma\gamma}^{\frac{3}{2}}}(Z_{\frac{\alpha}{2}}^2 - 1) = \pm Z_{\frac{\alpha}{2}}.$$

Thus, the lower confidence limit for $\gamma$is obtained by solving

$$T(\gamma) - \frac{K_3(T(\gamma))}{6I_{\gamma\gamma}^{\frac{3}{2}}}(Z_{\frac{\alpha}{2}}^2 - 1) = +Z_{\frac{\alpha}{2}}$$

and the upper confidence limit is obtained by solving

$$T(\gamma) - \frac{K_3(T(\gamma))}{6I_{\gamma\gamma}^{\frac{3}{2}}}(Z_{\frac{\alpha}{2}}^2 - 1) = -Z_{\frac{\alpha}{2}}.$$

These equations can be solved by using IMSL subroutine ZBREN. Denote the lower limit and upper limit of γ obtained by Bartlett's corrected procedure by $\hat{\gamma}_{BCcL}$ and $\hat{\gamma}_{BCcU}$. The corresponding lower limit and upper limit of the odds ratio are

$$\hat{\psi}_{BCcL} = e^{\hat{\gamma}_{BCcL}}$$

and

$$\hat{\psi}_{BCcU} = e^{\hat{\gamma}_{BCcU}}.$$

4.2 Confidence interval estimation based on the unconditional maximum likelihood estimator of common odds ratio.

4.2.1 Procedure based on the asymptotic properties of MLE

Gart(1962) showed that the asymptotic variance of the unconditional estimator $\hat{\psi}_u$ is

$$Var(\hat{\psi}_u) = \frac{\hat{\psi}^2}{V},$$

Where

$$V = \sum_{k=1}^{K} \hat{V}_k$$

and

$$(\hat{V}_k)^{-1} = (N_{1k}\hat{p}_{1k}\hat{q}_{1k})^{-1} + (N_{2k}\hat{p}_{2k}\hat{q}_{2k})^{-1}.$$

Also, from section 3.1,

$$p_{1k} = \frac{e^{\rho_k}}{1+e^{\rho_k}}, \qquad q_{1k} = \frac{1}{1+e^{\rho_k}}$$

and

$$p_{2k} = \frac{e^{\rho_k}}{e^{\gamma}+e^{\rho_k}}, \qquad q_{2k} = \frac{e^{\gamma}}{e^{\gamma}+e^{\rho_k}}.$$

Therefore,

$$(\hat{V}_k)^{-1} = \frac{N_{1k}(e^{\hat{\gamma}}+e^{\hat{\rho}_k})^2 + N_{2k}e^{\hat{\gamma}}(1+e^{\hat{\rho}_k})^2}{N_{1k}N_{2k}{}^{\hat{\gamma}}e^{\hat{\rho}_k}}.$$

An approximate $100(1-\alpha)\%$ confidence interval using the unconditional maximum likelihood estimator of the common odds ratio is given by

$$\hat{\psi}_u \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{\psi}^2}{\hat{V}}}.$$

Denote the lower limit and the upper limit of the unconditional odds ratio using asymptotic property of the mle by $\hat{\psi}_{MuL}$ and $\hat{\psi}_{MuU}$.

4.2.2 Procedure based on likelihood ratio

Let $\hat{p} = (\hat{p}_{11}, ..., \hat{p}_{1K})'$ and $\hat{\psi}$ be the maximum likelihood estimator of $p = (p_{11}, ..., p_{1K})'$ and ψ respectively and the corresponding maximized log-likelihood be $l(\hat{\psi},\hat{p})$. Further, for a given ψ, let $\tilde{p} = (\tilde{p}_{11}, ..., \tilde{p}_{1K})'$ be the maximum likelihood estimator of $p = (p_{11}, \cdots, p_{1K})'$and the corresponding log-likelihood by $l(\psi, \tilde{p})$. Then following the procedure discussed in chapter 2, the confidence interval using the likelihood procedure is obtained by solving

$$2\left[l(\hat{\psi},\hat{p}) - l(\psi,\tilde{p})\right] \leq \chi^2_{(1-\alpha)}(1).$$

From chapter 3, the maximized log-likelihood $l(\hat{\psi},\hat{p})$, using the parametrization of γ and ρ_k is

$$l = C + \sum_{k=1}^{K} \left(T_k\hat{\rho}_k + (N_{2k} - X_{2k})\hat{\gamma} - N_{1k}log(1 + e^{\hat{\rho}_k}) - N_{2k}log(e^{\hat{\gamma}} + e^{\hat{\rho}_k})\right).$$

We still need to find $l(\psi, \tilde{p})$. Again from section 3.2, we have

$$l(\psi,p) = C + \sum_{k=1}^{K} \left(X_{1k}log(\frac{p_{1k}}{q_{1k}}) + N_{1k}logq_{1k} + X_{2k}log\frac{p_{1k}}{\psi q_{1k}} + N_{2k}log\frac{\psi q_{1k}}{\psi q_{1k} + p_{1k}}\right).$$

For a given ψ, the maximum likelihood estimator for p_{1k}, $k = 1, \cdots, K$ is obtained by solving $\frac{\partial l}{\partial p_{1k}} = 0$. Now, the partial derivative of l with respect to p_{1k} is

$$\frac{\partial l}{\partial p_{1k}} = \frac{X_{1k}}{p_{1k}} - \frac{(N_{1k} - X_{1k})}{q_{1k}} + \frac{X_{2k}}{p_{1k}} - \frac{(N_{2k} - X_{2k})}{q_{1k}} - \frac{(N_{2k})(1-\psi)}{\psi q_{1k} + p_{1k}}$$

$$= \frac{X_{1k} + X_{2k}}{p_{1k}} - \frac{(N_{1k} + N_{2k} - X_{1k} - X_{2k})}{q_{1k}} - \frac{N_{2k}(1-\psi)}{\psi q_{1k} + p_{1k}} = 0.$$

From this and using the reparametrization of $\rho_k = log(\frac{p_{1k}}{1-p_{1k}})$ and $\gamma = log\psi$ we obtain

$$T_k\frac{(1 + e^{\rho_k})}{e^{\rho_k}} - \frac{(N_k - T_k)(1 + e^{\rho_k})}{1} - \frac{N_{2k}(1 + e^{\rho_k})(1 - e^{\gamma})}{e^{\gamma} + e^{\rho_k}} = 0,$$

which can be written as

$$A_k X_k^2 + B_k X_k + C_k = 0$$

where

$$B_k = -T_k(1+e^{\gamma}) + N_{1k}e^{\gamma} + N_{2k},$$

$$A_k = (N_k - T_k),$$

$$C_k = -T_k e^{\gamma}$$

and

$$X_k = e^{\rho_k}.$$

Now, $p_{1k} \epsilon (0,1)$, $\rho_k \epsilon (-\infty, \infty)$ and $e^{\rho_k} \epsilon (0, \infty)$. The solution of the quadratic equation is

$$X_k = \frac{-B_k \pm \sqrt{B_k^2 - 4A_k C_k}}{2A_k}.$$

We have to show that, it has two real roots and only one root is admissible. That is, only one solution is in the range $(0, \infty)$. Now $A_k = N_k - T_k$ is positive, $-C_k = T_k e^{\gamma}$ is positive therefore, $-A_K C_k$ is positive.

$$-A_k C_k > 0 \Rightarrow B_k^2 - 4A_k C_k > 0. \tag{4.6}$$

Therefore, the quadratic equation has two real roots. Now,

$$-4A_k C_k > 0 \Rightarrow B_k^2 - 4A_k C_k > B_k^2,$$

$$\sqrt{B_k^2 - 4A_k C_k} > B_k. \tag{4.7}$$

From this it is clear that $-B_k + \sqrt{B_k^2 - 4A_K C_k} > 0$ and $-B_k - \sqrt{B_k^2 - 4A_k C_k} < 0$. Therefore, we have only one admissible solution. Using this, the maximum likelihood estimate of e^{ρ_k} for a given e^{γ} is

$$e^{\tilde{\rho}_k} = \frac{-B_k + \sqrt{B_k^2 - 4A_k C_k}}{2A_k} \tag{4.8}$$

Putting this in $l(\gamma, \rho)$ we obtain

$$l(\gamma, \tilde{\rho}) = C + \sum_{k=1}^{K} \left(T_k \tilde{\rho}_k + (N_{2k} - X_{2k})\gamma - N_{1k} log(1 + e^{\tilde{\rho}_k}) - N_{2k} log(e^{\gamma} + e^{\tilde{\rho}_k}) \right).$$

From section 2.2, an approximate $100(1-\alpha)\%$ confidence interval for γ is obtained by solving

$$LR = 2\left(l(\hat{\gamma}, \hat{\rho}) - l(\gamma, \tilde{\rho})\right) \leq \chi^2_{1-\alpha}(1)$$

The above equation can be solved by using IMSL subroutine ZREAL. Denote the lower limit and upper limit of γ obtained by the likelihood procedure by $\hat{\gamma}_{LuL}$ and $\hat{\gamma}_{LuU}$. The corresponding lower limit and upper limit of the odds ratio are

$$\hat{\psi}_{LuL} = e^{\hat{\gamma}_{LuL}}$$

and

$$\psi_{LuU} = e^{\hat{\gamma}_{LuU}}.$$

4.2.3 Procedure based on adjusted likelihood ratio

In this approach (unconditional), γ is a scalar parameter and $\rho = (\rho_1, \cdots, \rho_K)'$ is the vector of nuisance parameters. As discussed in chapter 2, according to the procedure developed by Diciccio, Field and Fraser (1990), the pivotal statistics are

$$P_k = \frac{(\rho_k - \hat{\rho}_k)}{\hat{\psi}}, \qquad k = 1, \cdots, K$$

and

$$P_{K+1} = log\frac{\psi}{\hat{\psi}}.$$

Thus,

$$\rho_k = \hat{\rho}_k + P_k \hat{\psi}$$

and

$$\gamma = \hat{\gamma} + P_{K+1}.$$

From chapter 3, the log-likelihood can be written as

$$l = C + \sum_{k=1}^{K} (T_k \rho_k + (N_{2k} - X_{2k})\gamma - N_{1k} log(1 + e^{\rho_k}) - N_{2k} log(e^{\gamma} + e^{\rho_k})) .$$

Therefore, in terms of the pivotals, the log-likelihood is

$$l(P) = C + \sum_{k=1}^{K} \left(T_k(\hat{\rho}_k + P_k \hat{\psi}) + (N_{2k} - X_{2k})(\hat{\gamma} + P_{K+1}) \right) -$$

$$\sum_{k=1}^{K} \left(N_{1k} log(1 + e^{\hat{\rho}_k + P_k \hat{\psi}}) + N_{2k} log(e^{\hat{\gamma} + P_{K+1}} + e^{\hat{\rho}_k + P_k \hat{\psi}}) \right) \tag{4.9}$$

Hence,

$$l(0) = C + \sum_{k=1}^{K} (T_k \hat{\rho}_k + (N_{2k} - X_{2k})\hat{\gamma} - N_{1k} log(1 + e^{\hat{\rho}_k}) - N_{2k} log(e^{\hat{\gamma}} + e^{\hat{\rho}_k})) .$$

Now, we need to find $l(\tilde{P}(P_{K+1}))$. From equation (4.9) we obtain

$$\frac{\partial l}{\partial P_k} = T_k \hat{\psi} - \frac{N_{1k} e^{\hat{\rho}_k + P_k \hat{\psi}} (\hat{\psi})}{(1 + e^{\hat{\rho}_k + P_k \hat{\psi}})} - \frac{N_{2k} e^{\hat{\rho}_k + P_k \hat{\psi}} (\hat{\psi})}{e^{\hat{\gamma} + P_{K+1}} + e^{\hat{\rho}_k + P_k \hat{\psi}}} = 0$$

$$\Rightarrow T_k - \frac{N_{1k} e^{\hat{\rho}_k + P_k \hat{\psi}}}{(1 + e^{\hat{\rho}_k + P_k \hat{\psi}})} - \frac{N_{2k} e^{\hat{\rho}_k + P_k \hat{\psi}}}{e^{\hat{\gamma} + P_{K+1}} + e^{\hat{\rho}_k + P_k \hat{\psi}}} = 0$$

$$\Rightarrow T_k(1 + e^{\hat{\rho}_k + P_k \hat{\psi}})(e^{\hat{\gamma} + P_{K+1}} + e^{\hat{\rho}_k + P_k \hat{\psi}}) - N_{1k} e^{\hat{\rho}_k + P_k \hat{\psi}} (e^{\hat{\gamma} + P_{K+1}} + e^{\hat{\rho}_k + P_k \hat{\psi}})$$

$$-N_{2k} e^{\hat{\rho}_k + P_k \hat{\psi}} (1 + e^{\hat{\rho}_k + P_k \hat{\psi}}) = 0 \tag{4.10}$$

Let $X_k = e^{\hat{\psi} P_k}$. Then equation (4.10) can be written as

$$A_k X_k^2 + \mathrm{B}_k X_k + C_k = 0$$

where

$$A_k = (N_k - T_k) e^{2\hat{\rho}_k},$$

$$B_k = -\left(T_k(1 + e^{\hat{\gamma} + P_{K+1}}) - N_{1k} e^{\hat{\gamma} + P_{K+1}} - N_{2k} \right) e^{\hat{\rho}_k}$$

and

$$C_k = -T_k e^{\hat{\gamma}+P_{K+1}}.$$

Now, using the same argument as in the derivation in section 4.2.1, we have the values for P_k ($k = 1, \cdots, K$) that maximize $l(P)$ for a given value of P_{K+1} as

$$\tilde{P}_k = \frac{1}{\hat{\psi}} log\left(\frac{-B_k + \sqrt{B_k^2 - 4A_kC_k}}{2A_k}\right).$$

Substituting this value in equation (4.9), we obtain the maximized log-likelihood for a given value of P_{K+1} as

$$l(\tilde{P}(P_{K+1})) = C + \sum_{k=1}^{K}\Big(T_k(\hat{\rho}_k + \tilde{P}_k\hat{\psi}) + (N_{2k} - X_{2k})(\hat{\gamma} + P_{K+1})\Big) -$$
$$\sum_{k=1}^{K}\Big(N_{1k}log(1 + e^{\hat{\rho}_k+\tilde{P}_k\hat{\psi}}) + N_{2k}log(e^{\hat{\gamma}+P_{K+1}} + e^{\hat{\rho}_k+\tilde{P}_k\hat{\psi}})\Big) \qquad (4.11)$$

Therefore, the likelihood ratio statistic for the pivotal P_{K+1} is

$$LR_{K+1} = 2\left[l(0) - l(\tilde{P}(P_{K+1}))\right]$$
$$= 2\sum_{k=1}^{K}\Big(-T_k\tilde{P}_k\hat{\psi}) + (N_{2k} - X_{2k})(-P_{K+1})\Big)$$
$$-2\sum_{k=1}^{K}\left(N_{1k}log\frac{1 + e^{\hat{\rho}_k}}{1 + e^{\hat{\rho}_k+\tilde{P}_k\hat{\psi}}} + N_{2k}log\frac{e^{\hat{\gamma}} + e^{\hat{\rho}_k}}{e^{\hat{\gamma}+P_{K+1}} + e^{\hat{\rho}_k+\tilde{P}_k\hat{\psi}}}\right).$$

From chapter 2, the signed root statistic is

$$SR_{K+1} = -\sqrt{LR_{K+1}}, \qquad if \quad \gamma < \hat{\gamma}$$

and

$$SR_{K+1} = +\sqrt{LR_{K+1}}, \qquad if \quad \gamma > \hat{\gamma}.$$

From equation (4.9), we have

$$\frac{\partial l(P)}{\partial P_k} = T_k\hat{\psi} - \frac{N_{1k}\hat{\psi}e^{P_k\hat{\psi}+\hat{\rho}_k}}{1 + e^{\hat{\rho}_k+P_k\hat{\psi}}} - \frac{N_{2k}e^{\hat{\rho}_k+P_k\hat{\psi}}(\hat{\psi})}{e^{\hat{\gamma}+P_{K+1}} + e^{\hat{\rho}_k+P_k\hat{\psi}}}$$

and

$$\frac{\partial l(P)}{\partial P_{K+1}} = \sum_{k=1}^{K}(N_{2k} - X_{2k}) - \sum_{k=1}^{K} \frac{N_{2k}e^{\hat{\gamma}+P_{K+1}}}{e^{\hat{\gamma}+P_{K+1}} + e^{\hat{\rho}_k+P_k\hat{\psi}}}.$$

Hence,

$$-\frac{\partial^2 l(P)}{\partial P_k^2} = \frac{N_{1k}\hat{\psi}^2 e^{\hat{\rho}_k+P_k\hat{\psi}}}{(1+e^{\hat{\rho}_k+P_k\hat{\psi}})^2} + \frac{N_{2k}e^{\hat{\rho}_k+P_k\hat{\psi}}e^{\hat{\gamma}+P_{K+1}}(\hat{\psi}^2)}{(e^{\hat{\gamma}+P_{K+1}} + e^{\hat{\rho}_k+P_k\hat{\psi}})^2},$$

$$-\frac{\partial^2 l(P)}{\partial P_k \partial P_{K+1}} = -\frac{N_{2k}\hat{\psi}e^{\hat{\gamma}+P_{K+1}}e^{\hat{\rho}_k+P_k\hat{\psi}}}{(e^{\hat{\gamma}+P_{K+1}} + e^{\hat{\rho}_k+P_k\hat{\psi}})^2}$$

and

$$-\frac{\partial^2 l(P)}{\partial P_{K+1}^2} = \sum_{k=1}^{K} \frac{N_{2k}\hat{\psi}e^{\hat{\gamma}+P_{K+1}}e^{\hat{\rho}_k+P_k\hat{\psi}}}{(e^{\hat{\gamma}+P_{K+1}} + e^{\hat{\rho}_k+P_k\hat{\psi}})^2}$$

Therefore, $\frac{\partial^2 l(P)}{\partial P_k^2}$ for a given $P_k = 0$, $k = 1, \cdots, K+1$ is

$$-\frac{\partial^2 l(0)}{\partial P_k^2} = \frac{N_{1k}\hat{\psi}^2 e^{\hat{\rho}_k}}{(1+e^{\hat{\rho}_k})^2} + \frac{N_{2k}e^{\hat{\rho}_k}e^{\hat{\gamma}}(\hat{\psi}^2)}{(e^{\hat{\gamma}} + e^{\hat{\rho}_k})^2} = a_k$$

$$-\frac{\partial^2 l(0)}{\partial P_k \partial P_{K+1}} = -\frac{N_{2k}e^{\hat{\gamma}}e^{\hat{\rho}_k}(\hat{\psi})}{(e^{\hat{\gamma}} + e^{\hat{\rho}_k})^2} = b_k$$

and

$$-\frac{\partial^2 l(0)}{\partial P_{K+1}} = \sum_{k=1}^{K} \frac{N_{2k}e^{\hat{\gamma}}e^{\hat{\rho}_k}(\hat{\psi})}{(e^{\hat{\gamma}} + e^{\hat{\rho}_k})^2} = d.$$

The information matrix I can be written as

$$I = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix}, \quad |I| = |I_{11}|.|I_{22} - I_{21}I_{11}^{-1}I_{12}|$$

where I_{11} is a diagonal matrix with kth diagonal element a_k, I_{12} is a $K \times 1$ matrix with kth element b_k, $I_{12} = I_{21}$ and $I_{22} = d$ is a scalar.

Therefore, the determinant of the information matrix I (given $P_k = 0$ and $P_{K+1} = 0$) is

$$|I| = \left(\prod_{k=1}^{K} a_k\right)\left(d - \sum_{k=1}^{K} \frac{(b_k)^2}{a_k}\right)$$

and the determinant of the submatrix I^* for given $\tilde{P}_k$ $(k = 1, \cdots, K)$ is

$$|I^*| = \left(\prod_{k=1}^{K} - \frac{\partial^2 l\tilde{P}(P_{K+1})}{\partial P_k^2} \right).$$

$$= \prod_{k=1}^{K} \left[\frac{N_{1k}\hat{\psi}^2 e^{\hat{\rho}_k + \tilde{P}_k \hat{\psi}}}{(1 + e^{\hat{\rho}_k + \tilde{P}_k \hat{\psi}})^2} + \frac{N_{2k} e^{\hat{\rho}_k + \tilde{P}_k \hat{\psi}} e^{\hat{\gamma} + P_{K+1}} \hat{\psi}^2}{(e^{\hat{\gamma} + P_{K+1}} + e^{\hat{\rho}_k + \tilde{P}_k \hat{\psi}})^2} \right].$$

From the equation (4.9), we have,

$$l_1 \left(\tilde{P}(P_{K+1}) \right) = \left. \frac{\partial l(P)}{\partial P_{K+1}} \right|_{P_k = \tilde{P}_k} = \sum_{k=1}^{K} \left((N_{2k} - X_{2k}) - \frac{N_{2k} e^{\hat{\gamma} + P_{K+1}}}{e^{\hat{\gamma} + P_{K+1}} + e^{\hat{\rho}_k + \tilde{P}_k \hat{\psi}}} \right).$$

Following the procedure in section 2.3, the marginal tail probability for the pivotal P_{K+1} can be written as

$$Pr(P_{K+1} \leq p_{K+1}) = \Phi(SR_{K+1}) + \phi(SR_{K+1}) \left(\frac{1}{SR_{K+1}} + \frac{|I|^{\frac{1}{2}}}{l_1(\tilde{P}(P_{K+1})|I^*|^{\frac{1}{2}}} \right)$$

Hence, the $100(1 - \alpha)\%$ approximate lower and upper confidence limits can be obtained by solving

$$\Phi(SR_{K+1}) + \phi(SR_{K+1}) \left(\frac{1}{SR_{K+1}} + \frac{|I|^{\frac{1}{2}}}{l_1(\tilde{P}(P_{K+1}))|I^*|^{\frac{1}{2}}} \right) = \frac{\alpha}{2}$$

and

$$\Phi(SR_{K+1}) + \phi(SR_{K+1}) \left(\frac{1}{SR_{K+1}} + \frac{|I|^{\frac{1}{2}}}{l_1(\tilde{P}(P_{K+1}))|I^*|^{\frac{1}{2}}} \right) = 1 - \frac{\alpha}{2}$$

respectively. Denote these as P_L and P_U. Therefore, the corresponding lower and upper limits of γ using Diciccio's procedure are $\hat{\gamma} e^{P_L}$ and $\hat{\gamma} e^{P_U}$ respectively. Then the estimators for the lower and the upper limits of the confidence interval for ψ are

$$\hat{\psi}_{DuL} = e^{\hat{\gamma} e^{P_L}}$$

and

$$\hat{\psi}_{DuU} = e^{\hat{\gamma}_c^{P_U}}.$$

4.2.4 Bartlett's procedure based on the likelihood score

In this approach, we are interested to construct confidence interval for the parameter ψ, in the presence of nuisance parameters $\phi = (p_{11}, ..., p_{1K})'$. From section 2.3 (chapter 2), Bartlett's alternative to the maximum likelihood estimate $\hat{\psi}$,

$$T(\psi) = \frac{\partial l}{\partial \psi} - I_{\psi\phi} I_{\phi\phi}^{-1} (\frac{\partial l}{\partial \phi})$$

with variance $I_{\psi\psi.\phi}$ has asymptotically normal distribution. As reviewed in section 2.3, the $100(1-\alpha)\%$ confidence interval for ψ is obtained by solving

$$\frac{T(\psi)}{\sqrt{I_{\psi\psi.\phi}}} = \pm Z_{\frac{\alpha}{2}}$$

where $Z_{\frac{\alpha}{2}}$ is an approximate quantile of a standard normal distribution. From the unconditional log-likelihood discussed in chapter 3, the partial derivative of l with respect to ψ is

$$\frac{\partial l}{\partial \psi} = \sum_{k=1}^{K} \frac{-N_{2k} q_{1k}}{\psi q_{1k} + p_{1k}} + \sum_{k=1}^{K} \frac{(N_{2k} - X_{2k})}{\psi}.$$

The second partial derivative of l with respect to ψ is

$$\frac{\partial^2 l}{\partial \psi^2} = \sum_{k=1}^{K} \frac{N_{2k} q_{1k}^2}{(\psi q_{1k} + p_{1k})^2} - \sum_{k=1}^{K} \frac{(N_{2k} - X_{2k})}{\psi^2}.$$

Furthermore,

$$E(X_{2k}) = N_{2k} p_{2k}.$$

Therefore,

$$-E(\frac{\partial^2 l}{\partial \psi^2}) = -\sum_{k=1}^{K} \frac{N_{2k} q_{1k}^2}{(\psi q_{1k} + p_{1k})^2} + \sum_{k=1}^{K} \frac{N_{2k} q_{2k}}{\psi^2}.$$

Reparametrization of the above equation in terms of γ and ρ_k leads to

$$-E(\frac{\partial^2 l}{\partial \psi^2}) = \sum_{k=1}^{K} \frac{N_{2k} e^{\rho_k}}{(e^{\rho_k} + e^{\gamma})^2 e^{\gamma}}.$$

But from Fisher information matrix

$$I_{\psi\psi} = -E(\frac{\partial^2 l}{\partial \psi^2}).$$

Therefore,

$$I_{\psi\psi} = \sum_{k=1}^{K} \frac{N_{2k} e^{\rho_k}}{(e^{\rho_k} + e^{\gamma})^2 e^{\gamma}} = S(say). \qquad (4.12)$$

Furthermore, reparametrization of $\frac{\partial l}{\partial \psi}$ in terms of ρ_k and γ gives

$$\frac{\partial l}{\partial \psi} = \sum_{k=1}^{K} \frac{N_{2k} e^{\rho_k} - X_{2k}(e^{\gamma} + e^{\rho_k})}{e^{\gamma}(e^{\gamma} + e^{\rho_k})}. \qquad (4.13)$$

From the unconditional log-likelihood in chapter 3, we have

$$\frac{\partial l}{\partial p_{1k}} = \frac{(X_{1k} + X_{2k})}{p_{1k}} - \frac{(N_{1k} + N_{2k} - X_{2k} - X_{1k})}{q_{1k}} - \frac{N_{2k}(1-\psi)}{\psi q_{1k} + p_{1k}}.$$

But

$$\phi = (p_{11}, ..., p_{1K})'.$$

Therefore,

$$\frac{\partial^2 l}{\partial p_{1k} \partial p_{1k'}} = 0, \qquad k \neq k'.$$

Further, when $k = k'$

$$\frac{\partial^2 l}{\partial p_{1k}^2} = -\frac{X_{1k} + X_{2k}}{p_{1k}^2} - \frac{N_{2k} + N_{1k} - X_{1k} - X_{2k}}{q_{1k}^2} + \frac{N_{2k}(1-\psi)^2}{(\psi q_{1k} + p_{1k})^2}$$

and

$$-E(\frac{\partial^2 l}{\partial p_{1k}^2}) = \frac{N_{1k} p_{1k} + N_{2k} p_{2k}}{p_{1k}^2} + \frac{N_{2k} q_{2k} + N_{1k} q_{1k}}{q_{1k}^2} - \frac{N_{2k}(1-\psi)^2}{(\psi q_{1k} + p_{1k})^2}.$$

Using the reparametrization $p_{1k} = \frac{e^{\rho_k}}{1+e^{\rho_k}}$, we have

$$-E(\frac{\partial^2 l}{\partial p_{1k}^2}) = \frac{N_{1k}(1+e^{\rho_k})^2}{e^{\rho_k}} + \frac{N_{2k}(1+e^{\rho_k})^4}{e^{\rho_k}(e^{\gamma}+e^{\rho_k})^2}$$

and

$$-E(\frac{\partial^2 l}{\partial p_{1k}p_{1k'}}) = 0.$$

Therefore, the matrix $I_{\phi\phi}$ is diagonal, and is given by

$$I_{\phi\phi} = \begin{pmatrix} i_{11} & 0 & \cdots & 0 \\ 0 & i_{22} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & i_{KK} \end{pmatrix}$$

where

$$i_{kk} = \frac{N_{1k}(1+e^{\rho_k})^2}{e^{\rho_k}} + \frac{N_{2k}(1+e^{\rho_k})^4 e^{\gamma}}{e^{\rho_k}(e^{\rho_k}+e^{\gamma})^2}.$$

Now, from $\frac{\partial l}{\partial \psi}$, we have

$$\frac{\partial^2 l}{\partial p_{1k}\partial\psi} = \frac{N_{2k}(1-\psi)q_{1k}}{(\psi q_{1k}+p_{1k})^2} + \frac{N_{2k}}{\psi q_{1k}+p_{1k}},$$

$$= \frac{N_{2k}(q_{1k}-\psi q_{1k}+\psi q_{1k}+p_{1k})}{(\psi q_{1k}+p_{1k})^2},$$

$$= \frac{N_{2k}}{(\psi q_{1k}+p_{1k})^2}.$$

By reparametrization of $p_{1k} = \frac{e^{\rho_k}}{1+e^{\rho_k}}$, we have

$$\frac{\partial^2 l}{\partial p_{1k}\partial\psi} = \frac{N_{2k}(1+e^{\rho_k})^2}{(e^{\gamma}+e^{\rho_k})^2}.$$

But from Fisher information matrix,

$$I_{\phi\psi} = -E(\frac{\partial^2 l}{\partial\phi\partial\psi}).$$

Therefore,

$$I_{\phi\psi} = \begin{pmatrix} r_1 \\ \vdots \\ r_K \end{pmatrix}$$

with kth row element r_k, where

$$r_k = -\frac{N_{2k}(1+e^{\rho_k})^2}{(e^{\gamma}+e^{\rho_k})^2}.$$

Similarly,

$$I_{\psi\phi} = (\, r_1 \quad \cdots r_K \,).$$

As mentioned in chapter 2, we have

$$I_{\psi\psi.\phi} = I_{\psi\psi} - I_{\psi\phi} I_{\phi\phi}^{-1} I_{\phi\psi}.$$

But from the above derivations, $I_{\psi\phi} I_{\phi\phi}^{-1} I_{\phi\psi}$ is a scalar. Therefore,

$$I_{\psi\psi.\phi} = \sum_{k=1}^{K} \frac{N_{2k} e^{\rho_k}}{(e^{\rho_k}+e^{\gamma})^2 e^{\gamma}} - \sum_{k=1}^{K} \frac{r_k^2}{i_{kk}}. \tag{4.14}$$

Reparametrization of $\frac{\partial l}{\partial p_{1k}}$ in terms of γ and ρ_k gives

$$\frac{\partial l}{\partial p_{1k}} = \frac{T_k(1+e^{\rho_k})}{e^{\rho_k}} - (N_k - T_k)(1+e^{\rho_k}) - \frac{N_{2k}(1+e^{\rho_k})(1-e^{\gamma})}{(e^{\gamma}+e^{\rho_k})}$$

Therefore, $(\frac{\partial l}{\partial \phi})$ is a $K \times 1$ matrix and given by

$$(\frac{\partial l}{\partial \phi}) = \begin{pmatrix} u_1 \\ \vdots \\ u_K \end{pmatrix}$$

with kth row element u_k, which is equal to $\frac{\partial l}{\partial p_{1k}}$. From the above discussions, $I_{\psi\phi}I_{\phi\phi}^{-1}(\frac{\partial l}{\partial \phi})'$ is a scalar and is given by

$$I_{\psi\phi}I_{\phi\phi}^{-1}(\frac{\partial l}{\partial \phi}) = \sum_{k=1}^{K} r_k u_k (i_{kk})^{-1} \tag{4.15}$$

From equations (4.13), (4.14) and (4.15) the lower limit and the upper limit of γ are obtained by solving

$$\frac{T(\psi)}{\sqrt{I_{\psi\psi.\phi}}} = \pm Z_{\frac{\alpha}{2}} \tag{4.16}$$

that is by solving

$$\frac{\sum_{k=1}^{K}\left((D_k - r_k u_k (i_{kk})^{-1}\right)}{S} = \pm Z_{\frac{\alpha}{2}},$$

where

$$D_k = \sum_{k=1}^{K} \frac{N_{2k}e^{\rho_k} - X_{2k}(e^{\gamma} + e^{\rho_k})}{e^{\gamma}(e^{\gamma} + e^{\rho_k})},$$

$$r_k = -\frac{N_{2k}(1+e^{\rho_k})^2}{(e^{\gamma} + e^{\rho_k})^2},$$

$$u_k = \frac{T_k(1+e^{\rho_k})}{e^{\rho_k}} - (N_k - T_k)(1+e^{\rho_k}) - \frac{N_{2k}(1+e^{\rho_k})(1-e^{\gamma})}{(e^{\gamma}+e^{\rho_k})},$$

$$i_{kk} = \frac{N_{1k}(1+e^{\rho_k})^2}{e^{\rho_k}} + \frac{N_{2k}(1+e^{\rho_k})^4 e^{\gamma}}{e^{\rho_k}(e^{\rho_k}+e^{\gamma})^2},$$

and

$$S = \sum_{k=1}^{K} \frac{N_{2k}e^{\rho_k}}{(e^{\rho_k}+e^{\gamma})^2 e^{\gamma}}.$$

Note that the left hand side of the equation (4.16) involves e^{ρ_k}, which can be replaced from (4.8), by,

$$e^{\hat{\rho}_k} = \frac{-B_k + \sqrt{B_k^2 - 4A_kC_k}}{2A_k},$$

where

$$A_k = (N_k - T_k),$$

$$B_k = -T_k(1+e^{\gamma}) + N_{1k}e^{\gamma} + N_{2k},$$

and

$$C_k = -T_k e^{\gamma}.$$

Denote the lower limit and the upper limit of γ obtained by Bartlett's procedure by $\hat{\gamma}_{BucL}$ and $\hat{\gamma}_{BucU}$. The corresponding lower limit and upper limit of the odds ratio are

$$\hat{\psi}_{BucL} = e^{\hat{\gamma}_{BucL}}$$

and

$$\hat{\psi}_{BucU} = e^{\hat{\gamma}_{BucU}}.$$

4.2.5 Bartlett's procedure corrected for Bias and Skewness

In this procedure the nuisance parameters $\phi = (p_{11}, \cdots, p_{1K})'$ in $T(\psi)$ are replaced by their corresponding maximum likelihood estimates $\hat{\phi} = (\hat{p}_{11}, \cdots, \hat{p}_{1K})'$. This involves a bias of order $n^{\frac{-1}{2}}$. As reviewed in chapter 2, bias in $T(\psi)$ is given by

$$Bias(T_\psi) = -\frac{1}{2} trace\left(I_{\phi\phi}^{-1}\left(E(\frac{\partial^3 l}{\partial\psi\partial\phi\partial\phi^T}) + 2\frac{\partial I_{\psi\phi}}{\partial\phi}\right)\right)$$
$$+\frac{1}{2} trace\left(I_{\phi\phi}^{-1} M\right)$$

where M is the $K \times K$ array $(M_1, M_2, ...M_K)$ with j th column given by

$$M_j = \left(E(\frac{\partial^3 l}{\partial\phi_j\partial\phi\partial\phi^T}) + 2\frac{\partial I_{\phi\phi}}{\partial\phi_j}\right) I_{\phi\phi}^{-1} I_{\phi\psi}.$$

For convenience, we consider

$$Bias(T_\psi) = B_1 + B_2,$$

where

$$B_1 = -\frac{1}{2} trace\left(I_{\phi\phi}^{-1}\left(E(\frac{\partial^3 l}{\partial\psi\partial\phi\partial\phi^T}) + 2\frac{\partial I_{\psi\phi}}{\partial\phi}\right)\right)$$

and

$$B_2 = +\frac{1}{2} trace\left(I_{\phi\phi}^{-1} M\right).$$

Now, let

$$\frac{\partial^2 l}{\partial p_{1k} \partial p_{1k}} = j_k.$$

Then

$$\frac{\partial^2 l}{\partial \phi \partial \phi^T} = \begin{pmatrix} j_1 & 0 & \cdots & 0 \\ 0 & j_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & j_K \end{pmatrix}$$

with kth diagonal element $\frac{\partial^2 l}{\partial p_{1k}^2}$ and

$$E(\frac{\partial^3 l}{\partial \psi \partial \phi \partial \phi^T}) = \begin{pmatrix} t_1 & 0 & \cdots & 0 \\ 0 & t_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & t_K \end{pmatrix}$$

with kth row diagonal element E($\frac{\partial^3 l}{\partial \psi \partial p_{1k} \partial p_{1k}}$) $= t_k$. From section 4.2.3 (chapter 4), we have

$$\frac{\partial l}{\partial \psi} = \sum_{k=1}^{K} -\frac{N_{2k} q_{1k}}{\psi q_{1k} + p_{1k}} + \frac{(N_{2k} - X_{2k})}{\psi}.$$

Hence

$$\frac{\partial^2 l}{\partial p_{1k} \partial \psi} = \frac{N_{2k}}{(\psi q_{1k} + p_{1k})^2}$$

and

$$\frac{\partial^3 l}{\partial p_{1k} \partial p_{1k} \partial \psi} = -\frac{2N_{2k}(1-\psi)}{(\psi q_{1k} + p_{1k})^3}.$$

By the parametrization of $p_{1k} = \frac{1}{1+e^{\rho_k}}$, we have

$$E(\frac{\partial^3 l}{\partial p_{1k}\partial p_{1k}\partial\psi}) = -\frac{2N_{2k}(1-e^{\gamma})(1+e^{\rho_k})^3}{(e^{\gamma}+e^{\rho_k})^3}. \tag{4.17}$$

That is

$$t_k = -\frac{2N_{2k}(1-e^{\gamma})(1+e^{\rho_k})^3}{(e^{\gamma}+e^{\rho_k})^3}.$$

From section 4.2.3, we have

$$I_{\psi\phi} = (\ -E(\frac{\partial^2 l}{\partial\psi\partial p_{11}}),\quad \cdots,\quad -E(\frac{\partial^2 l}{\partial\psi\partial p1K}))$$

and

$$-E(\frac{\partial^2 l}{\partial p_{1k}\partial\psi}) = -\frac{N_{2k}}{(\psi q_{1k}+p_{1k})^2}.$$

Hence

$$\frac{\partial}{\partial p_{1k}}(-E(\frac{\partial^2 l}{\partial p_{1k}\partial\psi})) = \frac{2N_{2k}(1-\psi)}{(\psi q_{1k}+p_{1k})^3}.$$

By the parameterization of $p_{1k} = \frac{e^{\rho_k}}{1+e^{\rho_k}}$, we have

$$\frac{\partial}{\partial p_{1k}}(-E(\frac{\partial^2 l}{\partial p_{1k}\partial\psi})) = \frac{2N_{2k}(1-e^{\gamma})(1+e^{\rho_k})^3}{(e^{\gamma}+e^{\rho_k})^3}. \tag{4.18}$$

Therefore,

$$\frac{\partial I_{\psi\phi}}{\partial\phi} = \begin{pmatrix} l_1 & 0 & \cdots & 0 \\ 0 & l_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & l_K \end{pmatrix}$$

with kth diagonal element l_k, where

$$l_k = \frac{2N_{2k}(1-e^{\gamma})(1+e^{\rho_k})^3}{(e^{\gamma}+e^{\rho_k})^3}.$$

Therefore, from the above discussions $E(\frac{\partial^3 l}{\partial p_{1k}\partial p_{2k}\partial\psi})+2\frac{\partial I_{\psi\phi}}{\partial\phi}$ is a $K\times K$ matrix with diagonal element $t_k + 2l_k$. Also the diagonal element of the matrix $I_{\phi\phi}^{-1}$ is $(i_{kk})^{-1}$. Therefore,

$$B_1 = -\frac{1}{2}\sum_{k=1}^{K}\frac{(t_k + 2l_k)}{i_{kk}}. \tag{4.19}$$

Now, the jth array element of the matrix M is

$$M_j = \left(E(\frac{\partial^3 l}{\partial\phi_j \partial\phi\partial\phi^T}) + 2\frac{\partial I_{\phi\phi}}{\partial\phi_j}\right) I_{\phi\phi}^{-1} I_{\phi\psi}.$$

We have shown that

$$\frac{\partial^2 l}{\partial\phi\partial\phi^T} = \begin{pmatrix} j_1 & 0 & \cdots & 0 \\ 0 & j_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & j_K \end{pmatrix}.$$

Therefore,

$$\frac{\partial^3 l}{\partial\phi_k \partial\phi\partial\phi^T} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & s_k & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

where

$$s_k = \frac{\partial^3 l}{\partial p_{1k}^3}.$$

From section 4.2.3, we have

$$\frac{\partial^2 l}{\partial p_{1k}^2} = -\frac{(X_{1k} + X_{2k})}{p_{1k}^2} - \frac{(N_{1k} + N_{2k} - X_{1k} - X_{2k})}{q_{1k}^2}$$

$$+\frac{N_{2k}(1-\psi)^2}{(\psi q_{1k} + p_{1k})^2}.$$

Hence

$$\frac{\partial^3 l}{\partial p_{1k}^3} = \frac{2(X_{1k} + X_{2k})}{p_{1k}3} - \frac{2(N_{1k} + N_{2k} - X_{1k} - X_{2k})}{q_{1k}^3}$$

$$+\frac{N_{2k}(1-\psi)^3(-2)}{(\psi q_{1k}+p_{1k})^3}.$$

Therefore,

$$E(\frac{\partial^3 l}{\partial p_{1k}^3}) = \frac{2N_{1k}(q_{1k}-p_{1k})}{p_{1k}^2 q_{1k}^2} + \frac{2N_{2k}}{p_{1k}^2(\psi q_{1k}+p_{1k})}$$

$$-\frac{2N_{2k}(\psi)}{q_{1k}^2(\psi q_{1k}+p_{1k})} - \frac{2N_{2k}(1-\psi)^3}{(\psi q_{1k}+p_{1k})^3}.$$

By using the reparametrization of p_{1k}, q_{1k} and ψ, we have

$$E(\frac{\partial^3 l}{\partial p_{1k}^3}) = \frac{2N_{1k}(1-e^{\rho_k})(1+e^{\rho_k})^3}{(e^{\rho_k})^2} + \frac{2N_{2k}(1+e^{\rho_k})^3}{(e^{\rho_k})^2(e^{\gamma}+e^{\rho_k})}$$

$$-\frac{2N_{2k}(1+e^{\rho_k})^3 e^{\gamma}}{(e^{\gamma}+e^{\rho_k})} - \frac{2N_{1k}(1-e^{\gamma})^3(1+e^{\rho_k})^3}{(e^{\rho_k}+e^{\gamma})^3} = s_k \tag{4.20}$$

Let R.H.S of the above equation is equal to s_k.

We have already shown that $I_{\phi\phi}$ is a diagonal matrix with diagonal element i_{kk}.

Therefore,

$$\frac{\partial I_{\phi\phi}}{\partial \phi_k} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & v_k & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

where

$$v_k = \frac{\partial}{\partial p_{1k}} E(-\frac{\partial^2 l}{\partial p_{1k}^2}).$$

From section 4.2.4, we have

$$-E(\frac{\partial^2 l}{\partial p_{1k}^2}) = \frac{N_{1k}p_{1k}+N_{2k}p_{2k}}{p_{1k}^2} + \frac{N_{2k}q_{2k}+N_{1k}q_{1k}}{q_{1k}^2} - \frac{N_{2k}(1-\psi)^2}{(\psi q_{1k}+p_{1k})^2}.$$

Hence, using the reparametrization of p_{1k}, q_{1k} and ψ, we have

$$\frac{\partial}{\partial p_{1k}} E(-\frac{\partial^2 l}{\partial p_{1k}^2}) = \frac{N_{1k}(e^{\rho_k}-1)(1+e^{\rho_k})^3}{(e^{\rho_k})^2} - \frac{N_{2k}\left(e^{\gamma}(e^{\rho_k}-1)-2e^{\rho_k}\right)(1+e^{\rho_k})^3}{(e^{\rho_k})^2(e^{\gamma}+e^{\rho_k})^2}$$

$$+\frac{N_{2k}e^{\gamma}(e^{\rho_k}-1+2e^{\gamma})(1+e^{\rho_k})^3}{(e^{\gamma}+e^{\rho_k})^2}+\frac{2N_{2k}(1-e^{\gamma})^3(1+e^{\rho_k})^3}{(e^{\gamma}+e^{\rho_k})^3}=v_k \qquad (4.21)$$

Let the R.H.S of the above equation is equal to v_k But, kth array element of M is

$$M_k=\left(E(\frac{\partial^3 l}{\partial\phi_k\partial\phi\partial\phi^T})+2\frac{\partial I_{\phi\phi}}{\partial\phi_k}\right)I_{\phi\phi}^{-1}I_{\phi\psi}$$

Therefore,

$$M_k=\begin{pmatrix}0\\0\\m_k\\\vdots\\0\end{pmatrix},$$

with $m_k=(s_k+2v_k)(r_k)(i_{kk})^{-1}$. Therefore, M is a K×K matrix with diagonal element m_k. Hence,

$$I_{\phi\phi}^{-1}M=\begin{pmatrix}m_1i_{11}^{-1} & 0 & \cdots & 0\\ 0 & m_2i_{22}^{-1} & \cdots & 0\\ \vdots & & \ddots & \vdots\\ 0 & 0 & \cdots & m_Ki_{KK}^{-1}\end{pmatrix}.$$

Therefore,

$$B_2=\frac{1}{2}\sum_{k=1}^{K}\frac{(s_k+2v_k)(r_k)}{(i_{kk})^2}. \qquad (4.22)$$

From equations (4.19) and (4.22), we can find the bias in terms of γ and ρ_k. That is

$$Bias(T_\psi)=B_1+B_2=-\frac{1}{2}\sum_{k=1}^{K}\frac{(m_k+2l_k)}{i_{kk}}+\frac{1}{2}\sum_{k=1}^{K}\frac{(s_k+2v_k)(r_k)}{(i_{kk})^2}.$$

Bartlett's correction for Skewness

As reviewed in chapter 2, the third cumulant of T_ψ to the order $O(n^{\frac{-3}{2}})$ is obtained for $s,t,q=1,\cdots,K$ is given by

$$K_3(\psi)=2E(\frac{\partial^3 l}{\partial\psi^3})+3(\frac{\partial I_{\psi\psi}}{\partial\psi})$$

$$-3\sum_{s=1}^{K} f_s \left(2E(\frac{\partial^3 l}{\partial\psi^2\partial\phi_s}) + 2\frac{\partial I_{\psi\phi_s}}{\partial\psi} + \frac{\partial I_{\psi\psi}}{\partial\phi_s}\right)$$

$$+3\sum_s\sum_t f_s f_t \left(2E(\frac{\partial^3 l}{\partial\psi\partial\phi_s\partial\phi_t}) + \frac{\partial I_{\phi_s\phi_t}}{\partial\psi} + \frac{\partial I_{\psi\phi_t}}{\partial\phi_s} + \frac{\partial I_{\psi\phi_s}}{\partial\phi_t}\right)$$

$$-\sum_s\sum_t\sum_q f_s f_t f_q \left(2E(\frac{\partial^3 l}{\partial\phi_s\partial\phi_t\partial\phi_q}) + \frac{\partial I_{\phi_t}\phi_q}{\partial\phi_s} + \frac{\partial I_{\phi_s}\phi_q}{\partial\phi_t} + \frac{\partial I_{\phi_s}\phi_t}{\partial\phi_q}\right)$$

From section 4.2.4, we have

$$\frac{\partial^2 l}{\partial\psi^2} = \sum_{k=1}^{K} \frac{N_{2k}q_{1k}^2}{(\psi q_{1k}+p_{1k})^2} - \sum_{k=1}^{K}\frac{(N_{2k}-X_{2k})}{\psi^2}.$$

Hence

$$\frac{\partial^3 l}{\partial\psi^3} = \sum_{k=1}^{K} \frac{-2N_{2k}q_{1k}^2 q_{1k}}{(\psi q_{1k}+p_{1k})^3} + \sum_{k=1}^{K}\frac{2(N_{2k}-X_{2k})}{\psi^3}$$

and

$$E(\frac{\partial^3 l}{\partial\psi^3}) = \sum_{k=1}^{K} \frac{-2N_{2k}q_{1k}^3}{(\psi q_{1k}+p_{1k})^3} + \sum_{k=1}^{K}\frac{2N_{2k}q_{2k}}{\psi^3}. \quad (4.23)$$

Also from section 4.2.4, we have

$$I_{\psi\psi} = \sum_{k=1}^{K} \frac{-N_{2k}q_{1k}^2}{(\psi q_{1k}+p_{1k})^2} + \sum_{k=1}^{K}\frac{N_{2k}q_{2k}}{\psi^2}.$$

Therefore,

$$\frac{\partial I_{\psi\psi}}{\partial\psi} = \sum_{k=1}^{K} \frac{2N_{2k}q_{1k}^3}{(\psi q_{1k}+p_{1k})^3} - \sum_{k=1}^{K}\frac{N_{2k}q_{1k}(p_{1k}+2\psi q_{1k})}{\psi^2(p_{1k}+\psi q_{1k})}. \quad (4.24)$$

From equations (4.23) and (4.24), we have

$$2E(\frac{\partial^3 l}{\partial\psi^3}) + 3\frac{\partial I_{\psi\psi}}{\partial\psi} =$$

$$\sum_{k=1}^{K} \frac{2N_{2k}q_{1k}^3}{(\psi q_{1k}+p_{1k})^3} + \sum_{k=1}^{K} \frac{N_{2k}q_{1k}(p_{1k}-2\psi q_{1k})}{\psi^2(\psi q_{1k}+p_{1k})^2}. \tag{4.25}$$

Let the R.H.S of the above equation be A. From chapter 2 review, $f = I_{\phi\psi}I_{\phi\phi}^{-1}$, we have the sth element

$$f_s = \frac{(r_s)}{i_{ss}}.$$

$E(\frac{\partial^3 l}{\partial\psi^2\partial\phi_s})$ is a $K \times 1$ matrix and it's sth element can be determined as follows:

$$\frac{\partial^3 l}{\partial\psi^2\partial p_{1s}} = \frac{\partial}{\partial\psi}\left(\frac{\partial^2 l}{\partial\psi\partial p_{1s}}\right)$$

$$= \frac{\partial}{\partial\psi}\left(\frac{N_{2s}}{(\psi q_{1s}+p_{1s})^2}\right)$$

$$= \frac{-2N_{2s}q_{1s}}{(\psi q_{1s}+p_{1s})^3}$$

$$E(\frac{\partial^3 l}{\partial\psi^2\partial p_{1s}}) = \frac{-2N_{2s}q_{1s}}{(\psi q_{1s}+p_{1s})^3}.$$

$\frac{\partial I_{\psi\phi_s}}{\partial\psi}$ is a $K \times 1$ matrix and it's sth element can be determined as follows:

$$\frac{\partial I_{\psi\phi_s}}{\partial\psi} = \frac{\partial}{\partial\psi}\left(\frac{-N_{2s}}{(\psi q_{1s}+p_{1s})^2}\right)$$

$$= \frac{2N_{2s}q_{1s}}{(\psi q_{1s}+p_{1s})^3}$$

from the above equations, we have

$$2E(\frac{\partial^3 l}{\partial\psi^2\partial\phi_s}) + 2\frac{\partial I_{\psi\phi_s}}{\partial\psi} = 0.$$

From section 4.2.4, $I_{\psi\psi}$ is known and it is a scalar. Therefore, $\frac{\partial I_{\psi\psi}}{\partial\phi_s}$ is a $K \times 1$ matrix. It's sth element is

$$\frac{\partial I_{\psi\psi}}{\partial p_{1s}} = \frac{\partial}{\partial p_{1s}}\left(\sum_{s=1}^{K} \frac{-N_{2s}q_{1s}^2}{(\psi q_{1s}+p_{1s})^2} + \sum_{s=1}^{K} \frac{N_{2s}\psi q_{1s}}{\psi^2(\psi q_{1s}+p_{1s})}\right).$$

Hence,

$$\frac{\partial I_{\psi\psi}}{\partial p_{1s}} = \frac{2N_{2s}q_{1s}}{(\psi q_{1s} + p_{1s})^3} - \frac{N_{2s}}{\psi(\psi q_{1s} + p_{1s})^2}.$$

By reparametrization of p_{1s}, q_{1s} and ψ, we have

$$\frac{\partial I_{\psi\psi}}{\partial p_{1s}} = \frac{2N_{2s}(1 + e^{\rho_s})^2}{(e^{\gamma} + e^{\rho_s})^3} - \frac{N_{2s}(1 + e^{\rho_s})^2}{e^{\gamma}(e^{\gamma} + e^{\rho_s})^2}. \tag{4.26}$$

Let the R.H.S of the above equation be g_s. Therefore,

$$\sum_{s=1}^{K} f_s \left(2E(\frac{\partial^3 l}{\partial\psi^2 \partial\phi_s}) + 2\frac{\partial I_{\psi o_s}}{\partial\psi} + \frac{\partial I_{\psi\psi}}{\partial\phi_s}\right) = \sum_{s=1}^{K} \frac{(\tau_s)(g_s)}{i_{ss}}.$$

But from equation (4.17), the kth element of the matrix $E(\frac{\partial^3 l}{\partial\psi\partial o_s\partial o_t})$ is $\frac{-2N_{2k}(1-e^{\gamma})(1+e^{\rho_k})^3}{(e^{\gamma}+e^{\rho_k})^3}$ when $s = t = k$ and zero otherwise. We have already shown that the above value is t_k. From equation (4.18), the kth element of the matrix $\frac{\partial I_{\psi o_t}}{\partial o_s}$ is

$$= \frac{2N_{2k}(1 - e^{\rho_k})(1 + e^{\rho_k})^3}{(e^{\gamma} + e^{\rho_k})^3}$$

when $s = t = k$, otherwise zero and we have already shown that this value is equal to l_k.

Now, we need to calculate $\frac{\partial I_{o_s o_t}}{\partial\psi}$. When $s = t = k$, it is

$$\frac{\partial I_{o_k o_k}}{\partial\psi} = \frac{\partial}{\partial\psi}\left(\frac{N_{1k}}{p_{1k}q_{1k}} + \frac{N_{2k}p_{2k}}{p_{1k}^2} + \frac{N_{2k}q_{2k}}{q_{1k}^2} - \frac{N_{2k}(1-\psi)^2}{(\psi q_{1k} + p_{1k})^2}\right)$$
$$= \frac{N_{2k}(p_{1k} - q_{1k})}{p_{1k}q_{1k}(\psi q_{1k} + p_{1k})^2} + \frac{2N_{2k}(1-\psi)}{(\psi q_{1k} + p_{1k})^3}.$$

By reparametrization of p_{1k}, q_{1k} and ψ, we have

$$\frac{\partial I_{o_s o_t}}{\partial\psi} = \frac{N_{2k}(e^{\rho_k} - 1)(1 + e^{\rho_k})^3}{e^{\rho_k}(e^{\gamma} + e^{\gamma})^2} +$$

$$\frac{2N_{2k}(1 - e^{\gamma})(1 + e^{\rho_k})^3}{(e^{\gamma} + e^{\rho_k})^3} \tag{4.27}$$

when $s = t = k$. otherwise zero. Let the R.H.S of the above equation be e_k from eq 4.24 we have

$$2E(\frac{\partial^3 l}{\partial\psi^3}) + 3\frac{\partial I_{\psi\psi}}{\partial\psi} = A$$

and from equation (4.26). we have

$$-3\sum_{s=1}^{K} f_s \left(2E(\frac{\partial^3 l}{\partial\psi^2\partial\phi_s}) + 2\frac{\partial I_{\psi\phi_s}}{\partial\psi} + \frac{\partial I_{\psi\psi}}{\partial\phi_s}\right) =$$

$$-3\sum_{k=1}^{K}\frac{(r_k)(g_k)}{i_{kk}}.$$

From equation (4.27). we have

$$+3\sum_s\sum_t f_s f_t \left(2E(\frac{\partial^3 l}{\partial\psi\partial\phi_s\partial\phi_t}) + \frac{\partial I_{\phi_s\phi_t}}{\partial\psi} + \frac{\partial I_{\psi\phi_t}}{\partial\phi_s} + \frac{\partial I_{\psi\phi_s}}{\partial\phi_t}\right)$$

$$= 3\sum_{k=1}^{K}(\frac{r_k}{i_{kk}})^2(e_k).$$

From equation (4.20) and from equation (4.21). we have

$$-\sum_s\sum_t\sum_q f_s f_t f_q \left(2E(\frac{\partial^3 l}{\partial\phi_s\partial\phi_t\partial\phi_q}) + \frac{\partial I_{\phi_t\phi_q}}{\partial\phi_s} + \frac{\partial I_{\phi_s\phi_q}}{\partial\phi_t} + \frac{\partial I_{\phi_s\phi_t}}{\partial\phi_q}\right)$$

$$= -\sum_{k=1}^{K}(\frac{r_k}{i_{kk}})^3(2s_k + 3v_k)$$

Thus.

$$K_3(\psi) = A - 3\sum_{k=1}^{K}\frac{(r_k)(g_k)}{i_{kk}} + 3\sum_{k=1}^{K}(\frac{r_k}{i_{kk}})^2(e_k) - \sum_{k=1}^{K}(\frac{r_k}{i_{kk}})^3(2s_k + 3v_k).$$

The lower and upper confidence limits of γ are obtained by solving

$$\frac{T_\psi}{\sqrt{I_{\psi\psi.\phi}}} - \frac{B(T_\psi)}{\sqrt{I_{\psi\psi.\phi}}} - \frac{K_3(\psi)(Z^2_{\frac{\alpha}{2}} - 1)}{6(I_{\psi\psi.\phi})} = \pm Z_{\frac{\alpha}{2}} \tag{4.28}$$

Note that the left hand side of the equation (4.28) involves e^{ρ_k}, which can be replaced from (4.8), by,

$$e^{\tilde{\rho}_k} = \frac{-B_k + \sqrt{B_k^2 - 4A_kC_k}}{2A_k},$$

where

$$A_k = (N_k - T_k),$$

$$B_k = -T_k(1 + e^\gamma) + N_{1k}e^\gamma + N_{2k},$$

and

$$C_k = -T_ke^\gamma.$$

Denote the lower limit and the upper limit of γ obtained by Bartlett's corrected procedure by $\hat{\gamma}_{BCucL}$ and $\hat{\gamma}_{BCucU}$. The corresponding lower limit and upper limit of the odds ratio are

$$\hat{\psi}_{BCucL} = e^{\hat{\gamma}_{BCucL}}$$

and

$$\hat{\psi}_{BCucU} = e^{\hat{\gamma}_{BCucU}}.$$

CHAPTER 5

SIMULATION STUDIES

In this chapter, the performance of the likelihood based procedures except the adjusted likelihood procedure based on the unconditional likelihood derived in chapter 4. are examined through simulations. The adjusted likelihood based procedure based on the unconditional likelihood is showing some convergence problems that could not be resolved by the author. IMSL random number generator RNBIN was used to generate binomial variables. The range of values for the parameters K, N_{1k}, N_{2k}, ψ and p_{1k} used in the simulation studies were chosen to be representative of situations which arise in epidemiologic practice. In the simulation study, for each of the K =5, 10 strata, the sample sizes chosen were (N_{1k},N_{2k}) = (5,5), (10,10), (20,20), (5,20). The values of probabilities p_{1k} chosen were $p_{1k} = 0.05 + 0.04k(\frac{20}{K})$ [Robins, Breslow and Greenland (1986)] and the values of ψ chosen were ψ= 1, 3.5 and 6.5. For all the likelihood procedures and for each combination of K, (N_{1k}, N_{2k}), p_{1k} and ψ, we produced the tail and the coverage probabilities and the average lengths based on 1000 samples. The validity of the confidence interval is determined by the probability that the random interval covers the parameter value. This probability is called coverage probability. We have use 95% nominal confidence coefficient. The tail probabilities are the probabilities that the parameter value lies outside the random interval. Using the conventional rule, we added 0.5 to each observed frequency in any simulated table where a zero observed frequency occured. Tables 5.1a and 5.1b list the lower and upper tail probabilities, coverage probabilities and average length of the confidence intervals for the common odds ratio using the conditional likelihood. Tables 5.2a and 5.2b list the lower and upper tail probabilities, coverage probabilities and average length of the confidence intervals for the common odds ratio using the unconditional likelihood.

Results: The ML method (Procedure based on maximum likelihood estimate) provides adequate coverage ($p \geq 0.9$, where p is estimated coverage probability) for ψ= 1.0 and unacceptable coverage for other values of ψ for both conditional and unconditional likelihoods. The LR method (Procedure based on likelihood ratio) provides excellent coverage for $\psi = 1.0, 3.5$ and 6.5 for all designs used for conditional likelihood except for the design $N_{1k} = N_{2k} = 5$, $K = 10$ and $\psi = 6.5$. The LR method based on the unconditional likelihood also provides excellent coverage for all designs used. The methods B and BC (Bartlett and Bartlett's corrected) provide excellent coverage for ψ=1.0, 3.5 and 6.5 for all designs used for both conditional and unconditional likelihood except for the design $N_{1k} = N_{2k} = 5$,$K = 10$ and $\psi = 6.5$ for conditional likelihood. The SQ method (Signed square root of the likelihood ratio) provides excellent coverage ($p \geq$0.94) for ψ = 1,3.5 and 6.5 for the design $N_{1k} = N_{2k} = 20$ and $K = 5$ and also for $\psi = 3.5$ and 6.5 for the design $N_{1k} = N_{2k} = 10$ and $K = 10$. For all other designs the coverage dropped below 94% for conditional likelihood. The SQ method provides excellent coverage for the unconditional likelihood for all the designs used. From Tables 5.1a and 5.1b, we note that the likelihood ratio intervals provide the upper tail probabilities which are larger than those of the lower tail probabilities for many of the designs used for conditional likelihood. The SQ method gives higher values for lower tail probabilities than LR method for all designs used for conditional likelihood and most of the designs used for unconditional likelihood. For $N_{1k} = N_{2k} = 20$ and $K = 5$, the methods LR, B, BC and SQ performed equally well in terms of coverage and tail probabilities for conditional likelihood. But for the same design, when $\psi = 6.5$ these methods provide excellent upper tail probabilities and an adequate lower tail probabilities. For $N_{1k} = N_{2k} = 10$ and $K = 5$, the methods B and BC performed well in terms of tail and coverage probabilities when $\psi = 1.0$, for conditional likelihood. For $N_{1k} = N_{2k} = 20$ and $K = 10$, the SQ method performed well in terms of tail and coverage probabilities, when $\psi = 1.0$ and 3.5 for conditional likelihood. For $N_{1k} = N_{2k} = 10$ and $K = 10$, the methods B and BC performed well in terms of tail and coverage probabilities when $\psi = 1.0$ for conditional likelihood. For unconditional likelihood the SQ method (Signed square root of likelihood ratio) performed well in terms of coverage and tail probabilities when $K = 5$ for all the designs used. For $N_{1k} = N_{2k} = 20$ and $K = 10$, the methods B and SQ performed equally well in terms of tail and coverage probabilities for unconditional likelihood. For $N_{1k} = N_{2k} = 10$ and $K = 10$, the methods LR, B, BC and SQ performed well in terms of tail and coverage probabilities when $\psi = 1.0$ for unconditional likelihood. But for other values of ψ the methods LR, B, and SQ provide excellent upper tail probabilities and unacceptable lower tail probabilities. For the design $N_{1k} = 5$ and $N_{2k} = 20$ and $K = 5$, the methods LR, B, and BC performed well in terms of tail and coverage probabilities for $\psi = 1$ and 3.5.

In summary, for conditional likelihood, in terms of coverage probabilities, the methods LR, B and BC provide excellent coverage for all the designs used. But in terms of tail probabilities Bartlett's method perfomed slightly better than other method. For unconditional likelihood, in terms of coverage probabilities, the methods LR, B, BC and SQ provide excellent coverage for all the designs used. But in terms of tail probabilities, the methods B and SQ peformed well. However, in terms of average length the method B gave the shortest average length. For the Bartlett method, the unconditional likelihood gave the coverage probability closed to 0.95 and the tail probabilities closed to 0.25 for most of the designs used. Based on the results of these likelihood based procedures, the Bartletts method B with unconditional likelihood seems to be most suitable for constructing confidence limits for common odds ratio, atleast for the kinds of designs that have been used in the simulations study in this thesis. However, most of these procedures fall short of producing adequate coverage probability when K (≥ 25) increases and the sample sizes (< 5) in each tables are small. The likelihood procedure corrected for appropriate tails in small samples developed following Deciccio, Field and Fraser (1990) is expected to perform well in these situations.

Table 5.1a: Lower and upper tail probabilities, coverage probabilities and average lengths of the confidence intervals for the common odds ratio using the conditional likelihood.

p_{1k} = 0.05+0.04k(20/K), K=5, alpha=0.05

N_{1k}	N_{2k}	Psi	Method	Lower	Coverage	Upper	Length
20	20	1.0	ML	7.3	92.7	0.0	1.3
			LR	2.5	95.1	2.4	1.4
			B	2.6	95.0	2.4	1.0
			BC	2.6	95.0	2.4	1.4
			SQ	2.6	94.3	3.1	1.4
			DA	3.8	93.7	2.5	1.4
		3.5	ML	28.5	46.3	39.5	1.3
			LR	2.3	95.6	2.0	5.7
			B	1.8	96.2	2.8	5.5
			BC	2.0	96.0	2.1	5.7
			SQ	2.7	93.8	3.5	5.4
			DA	3.8	94.2	2.0	5.0
		6.5	ML	36.5	24.0	39.5	1.5
			LR	1.4	96.0	2.0	12.0
			B	1.3	95.8	2.8	11.4
			BC	1.5	96.4	2.1	12.0
			SQ	1.2	95.8	3.0	12.6
			DA	3.7	94.3	2.0	12.3
10	10	1.0	ML	8.7	91.3	0.0	1.8
			LR	1.4	96.4	2.4	2.2
			B	2.1	95.7	2.2	2.2
			BC	2.1	95.7	2.2	2.3
			SQ	2.2	93.6	4.2	2.2
			DA	5.8	90.0	4.2	2.2
		3.5	ML	26.8	47.3	25.9	1.9
			LR	0.9	97.0	2.1	8.9

N_{1k}	N_{2k}	Psi	Method	Lower	Coverage	Upper	Length
			B	1.0	96.4	2.6	8.3
			BC	1.3	96.4	2.3	9.0
			SQ	1.3	96.0	2.7	9.0
			DA	10.5	86.9	2.7	8.7
		6.5	ML	29.4	24.9	45.7	2.0
			LR	0.2	96.4	3.4	17.1
			B	0.6	95.5	3.9	15.8
			BC	0.9	95.5	3.6	17.0
			SQ	1.0	95.3	3.7	18.1
			DA	6.7	89.1	3.7	17.5
5	5	1.0	ML	0.0	91.9	45.7	2.0
			LR	2.0	97.9	0.1	3.9
			B	2.3	96.1	1.6	3.4
			BC	2.3	96.1	1.6	3.7
			SQ	2.5	91.3	6.1	3.4
			DA	3.4	91.9	4.7	3.5
		3.5	ML	0.0	87.2	12.3	8.5
			LR	0.2	99.5	0.3	12.7
			B	0.1	95.7	4.2	10.4
			BC	0.2	96.2	3.5	12.2
			SQ	0.3	93.4	6.2	11.7
			DA	1.2	93.1	5.7	12.9
		6.5	ML	0.0	81.5	18.4	13.6
			LR	0.0	99.6	0.4	22.5
			B	0.0	94.0	6.0	16.9
			BC	0.0	94.9	5.1	20.6
			SQ	0.0	93.2	6.8	20.2
			DA	1.9	91.8	6.3	23.1
5	20	1.0	ML	0.0	92.7	7.3	2.2

N_{1k}	N_{2k}	Psi	Method	Lower	Coverage	Upper	Length
			LR	2.5	97.2	0.3	2.8
			B	2.4	95.7	1.9	2.7
			BC	2.4	95.7	1.9	2.7
			SQ	2.6	90.2	7.2	2.1
			DA	4.0	90.1	5.9	2.4
		3.5	ML	0.0	91.5	8.5	8.1
			LR	1.3	97.3	1.4	10.0
			B	1.3	96.0	2.7	9.3
			BC	1.4	96.1	2.5	11.1
			SQ	2.6	90.5	6.8	15.1
			DA	3.5	91.2	5.3	9.5
		6.5	ML	0.0	90.4	9.5	15.1
			LR	0.4	97.3	2.3	19.3
			B	1.1	95.5	3.4	17.6
			BC	1.3	95.8	2.9	19.6
			SQ	1.5	92.3	6.2	17.4
			DA	2.9	92.5	4.6	19.4

Table 5.1b: Lower and upper tail probabilities, coverage probabilities and average lengths of the confidence intervals for the common odds ratio using the conditional likelihood.

p_{1k} = 0.05+0.04k(20/K), K=10, alpha=0.05

N_{1k}	N_{2k}	Psi	Method	Lower	Coverage	Upper	Length
20	20	1.0	ML	5.2	94.5	0.3	0.9
			LR	1.6	96.0	2.4	0.9
			B	1.7	95.9	2.4	0.9
			BC	1.7	95.9	2.4	0.9
			SQ	2.2	95.4	2.4	0.8
			DA	7.8	81.0	11.2	0.8
		3.5	ML	20.1	42.5	28.4	1.0
			LR	1.7	95.8	2.5	3.7
			B	1.9	95.5	2.6	3.6
			BC	1.9	95.6	2.5	3.7
			SQ	1.9	95.6	2.6	3.6
			DA	6.2	91.3	2.5	3.4
		6.5	ML	33.7	23.9	42.4	1.1
			LR	1.2	96.6	2.2	7.5
			B	1.1	96.3	2.6	7.4
			BC	1.1	96.4	2.5	7.4
			SQ	1.4	96.4	2.2	7.5
			DA	4.7	92.4	2.9	7.9
10	10	1.0	ML	6.1	93.9	0.0	1.2
			LR	2.4	95.0	2.6	1.4
			B	2.3	95.1	2.6	1.4
			BC	2.3	95.1	2.6	1.4
			SQ	2.8	92.0	5.2	1.3
			DA	5.8	89.5	4.2	1.1
		3.5	ML	23.0	44.7	32.3	1.3
			LR	0.7	96.1	3.2	5.2

N_{1k}	N_{2k}	Psi	Method	Lower	Coverage	Upper	Length
			B	0.7	96.1	3.2	5.0
			BC	0.7	96.1	3.2	5.2
			SQ	0.8	95.3	3.9	5.9
			DA	10.1	86.0	3.9	4.0
		6.5	ML	23.8	23.5	52.7	1.4
			LR	0.3	95.5	4.2	10.0
			B	0.2	94.9	4.9	9.6
			BC	0.8	95.2	4.5	10.0
			SQ	0.3	94.7	5.0	9.8
			DA	8.6	87.0	4.4	9.8
5	5	1.0	ML	0.0	93.5	6.5	1.8
			LR	1.2	97.3	1.5	2.1
			B	1.2	97.3	1.5	1.9
			BC	1.2	97.3	1.5	1.9
			SQ	3.2	84.7	12.1	1.5
			DA	5.6	87.8	6.6	1.7
		3.5	ML	0.0	86.3	13.7	5.4
			LR	0.9	96.6	3.3	6.5
			B	0.1	95.3	4.6	6.0
			BC	0.2	96.2	3.6	6.4
			SQ	0.3	89.4	10.3	5.6
			DA	0.8	90.9	8.3	5.9
		6.5	ML	0.0	74.6	25.4	8.5
			LR	0.0	92.1	7.9	10.4
			B	0.0	90.3	9.6	9.6
			BC	0.0	91.4	8.3	10.5
			SQ	0.0	86.1	13.9	9.5
			DA	1.2	87.3	11.5	9.9
5	20	1.0	ML	0.0	95.5	5.0	1.5

N_{1k}	N_{2k}	Psi	Method	Lower	Coverage	Upper	Length
			LR	1.8	96.4	1.0	1.6
			B	2.0	96.3	1.7	1.6
			BC	2.0	96.3	1.7	1.6
			SQ	8.2	68.2	23.5	0.9
			DA	5.1	83.7	11.2	1.2
		3.5	ML	0.0	94.5	5.5	5.2
			LR	0.8	96.7	2.5	5.7
			B	1.1	96.4	2.3	5.6
			BC	1.1	96.4	2.3	5.6
			SQ	3.0	82.2	14.8	4.1
			DA	3.1	90.1	6.8	4.6
		6.5	ML	0.0	92.9	7.1	9.6
			LR	0.2	96.8	3.0	10.8
			B	0.4	96.3	3.2	10.4
			BC	0.6	96.4	3.0	10.8
			SQ	0.2	88.3	10.9	8.7
			DA	1.4	90.9	7.7	9.3

Table 5.2a: Lower and upper tail probabilities, coverage probabilities and average lengths of the confidence intervals for the common odds ratio using the unconditional likelihood.
p_{1k} = 0.05+0.04k(20/K), K=5, alpha=0.05

N_{1k}	N_{2k}	Psi	Method	Lower	Coverage	Upper	Length
20	20	1.0	ML	6.6	93.4	0.0	1.3
			LR	2.7	95.0	2.3	1.5
			B	2.7	94.9	2.4	1.4
			BC	3.4	94.3	2.3	1.5
			SQ	2.8	94.3	2.3	1.5
		3.5	ML	24.2	56.9	18.9	1.9
			LR	2.5	95.4	2.0	6.1
			B	2.7	95.3	2.0	5.8
			BC	2.3	95.7	2.4	6.2
			SQ	3.0	95.0	2.0	6.1
		6.5	ML	31.4	44.2	24.4	2.9
			LR	1.5	97.0	1.5	13.6
			B	1.8	96.5	1.7	12.3
			BC	1.7	95.9	2.4	12.3
			SQ	2.2	96.2	1.5	12.5
10	10	1.0	ML	8.3	91.7	0.0	1.9
			LR	2.9	94.7	2.7	2.4
			B	2.7	94.4	2.8	2.3
			BC	3.1	94.8	2.1	2.4
			SQ	2.8	94.3	2.8	2.4
		3.5	ML	23.6	58.0	18.4	2.8
			LR	1.2	96.8	2.0	10.5
			B	1.6	96.3	2.1	9.4
			BC	1.4	95.7	2.9	9.8
			SQ	2.0	96.0	1.9	10.3
		6.5	ML	26.3	46.6	27.1	4.1

N_{1K}	N_{2K}	Psi	Method	Lower	Coverage	Upper	Length
			LR	0.1	97.0	2.9	21.7
			B	1.0	96.0	3.0	18.6
			BC	0.1	95.5	3.6	18.6
			SQ	1.3	96.1	2.5	21.4
5	5	1.0	ML	0.0	96.1	3.9	2.3
			LR	2.1	96.3	1.6	4.3
			B	2.5	95.8	1.7	3.8
			BC	3.7	94.4	1.9	4.1
			SQ	2.8	94.3	2.9	4.2
		3.5	ML	0.0	91.1	8.9	10.4
			LR	0.4	96.8	2.8	16.1
			B	0.3	96.1	3.5	13.4
			BC	0.6	94.9	4.5	13.6
			SQ	2.0	96.1	1.9	14.2
		6.5	ML	0.0	92.9	7.1	27.6
			LR	0.6	97.4	2.6	23.2
			B	0.1	95.9	4.0	22.9
			BC	0.1	94.0	5.9	24.9
			SQ	1.3	96.1	2.6	23.4
5	20	1.0	ML	0.0	92.7	7.3	2.2
			LR	2.0	95.9	2.1	2.8
			B	2.4	95.3	2.3	2.7
			BC	3.5	94.7	1.8	3.1
			SQ	2.4	95.1	2.5	2.8
		3.5	ML	0.0	91.5	8.5	8.1
			LR	1.7	96.3	2.0	11.6
			B	2.3	95.2	2.5	10.4
			BC	1.3	96.5	2.2	11.1
			SQ	3.1	94.8	2.1	11.5

N_{1k}	N_{2k}	Psi	Method	Lower	Coverage	Upper	Length
		6.5	ML	0.0	90.4	9.5	15.1
			LR	0.2	97.6	2.2	23.8
			B	1.7	95.9	2.4	20.9
			BC	0.3	96.2	3.5	22.7
			SQ	2.2	95.6	2.2	23.7

Table 5.2b: Lower and upper tail probabilities, coverage probabilities and average lengths of the confidence intervals for the common odds ratio using the unconditional likelihood.
p_{1k} = 0.05+0.04k(20/K), K=10, alpha=0.05

N_{1k}	N_{2k}	Psi	Method	Lower	Coverage	Upper	Length
20	20	1.0	ML	0.2	95.3	4.5	0.9
			LR	1.7	95.8	2.5	0.9
			B	1.8	95.7	2.5	1.0
			BC	2.0	95.3	2.7	1.0
			SQ	1.8	95.7	2.7	0.9
		3.5	ML	24.8	55.6	18.6	1.4
			LR	2.4	95.8	1.8	3.9
			B	2.8	94.9	2.3	3.8
			BC	1.9	95.1	3.0	3.9
			SQ	2.8	95.8	1.8	3.8
		6.5	ML	29.0	44.5	26.5	2.1
			LR	1.5	96.9	1.6	8.4
			B	2.1	96.2	1.7	7.8
			BC	1.2	96.2	2.6	8.1
			SQ	2.3	96.1	1.5	8.0
10	10	1.0	ML	0.0	93.9	6.1	1.4
			LR	2.7	94.4	2.9	1.5
			B	2.7	94.6	2.7	1.4
			BC	2.4	94.5	3.1	1.5
			SQ	2.7	94.4	2.9	1.5
		3.5	ML	0.0	96.1	3.8	7.9
			LR	1.0	96.5	2.5	5.8
			B	1.2	96.1	2.7	5.6
			BC	1.0	95.3	3.7	5.8
			SQ	1.3	96.2	2.5	5.8
		6.5	ML	0.0	97.2	2.8	20.6

N_{1K}	N_{2K}	Psi	Method	Lower	Coverage	Upper	Length
			LR	0.3	97.5	2.1	11.9
			B	0.4	97.2	2.4	11.0
			BC	0.3	95.3	4.4	11.8
			SQ	0.8	99.0	0.2	11.6
5	5	1.0	ML	0.1	96.7	3.2	1.5
			LR	1.6	95.9	2.5	2.3
			B	1.6	96.0	2.5	2.2
			BC	2.0	95.5	2.5	2.5
			SQ	1.9	95.6	2.5	2.2
		3.5	ML	0.0	94.9	5.1	6.3
			LR	0.6	97.3	2.1	7.9
			B	0.5	97.1	2.4	7.4
			BC	0.3	94.6	5.1	8.2
			SQ	0.7	97.1	2.2	7.9
		6.5	ML	0.0	95.4	4.6	16.4
			LR	0.6	96.1	3.9	12.3
			B	0.0	94.8	5.2	12.3
			BC	0.0	92.8	7.9	14.3
			SQ	0.0	96.0	4.0	13.4
5	20	1.0	ML	0.0	95.5	5.0	1.5
			LR	2.3	95.6	2.1	1.6
			B	2.0	96.1	1.9	1.6
			BC	2.1	95.7	2.2	1.5
			SQ	2.3	95.6	2.1	1.6
		3.5	ML	0.0	94.5	5.5	5.2
			LR	2.1	95.6	1.7	6.4
			B	2.4	95.6	1.7	6.4
			BC	2.4	95.6	2.0	6.2
			SQ	0.6	94.2	5.2	6.7

N_{1k}	N_{2k}	Psi	Method	Lower	Coverage	Upper	Length
		6.5	ML	0.0	92.9	7.1	9.6
			LR	1.5	96.8	1.7	13.3
			B	1.4	96.5	2.0	12.0
			BC	0.2	92.5	7.2	10.7
			SQ	1.6	96.7	1.7	12.6

REFERENCES

Bartlett, M. S. (1953). Approximate confidence intervals. I. Biometrika, 40, 12-19.

Breslow, N. E. and Liang, K. Y. (1982). The variance of the Mantel-Haenzel estimator. Biometrics, 38, 943-952.

Brown, C. C. (1981). The validity of approximation methods for interval estimation of the odds ratio. Amer. J. Epideiol., 113, 474-480.

Cochran, W. G. (1954). Some methods for strengthening the common χ^2 tests. Biometrics, 10, 417-451.

Deciccio, T. J. (1988). Likelihood inference for linear regression models. Biometrika, 75, 29-34.

Deciccio, T. J., Field, C. A. and Fraser, D. A. (1990). Approximation of marginal tail probabilities and inference for scaler parameter. Biometrica, 77, 77-96.

Fisher, R. A. (1935). The logic of inductive inference. J. Roy. Stat. Soc. A 98, 39-54.

Fraser, D. A. S. (1991). Statiatical Inference: Likelihood to Significance. J. Am. Statist. Assoc., 86, 258-265.

Gart, J. J. (1962). On the combination of relative risks. Biometrics, 18, 601-610.

Gart, J. J. (1970). Point and interval estimation of the common odds ratio in the combination of 2X2 tables with fixed marginals. Biometrica, 57, 471-475.

Harkness, W. L. (1965). Properties of the extended hypergeometric distribution. Ann. Math. Stat., 36, 938-945.

Hauck, W. W. (1984). A comparative study of conditional maximum likelihood estimation of a common odds ratio. Biometrics. 40, 1117-1123.

Hauck, W. W., Anderson, S. and Leahy, F. J. (1982). Finite-sample properties of some old and some new estimators of a common odds ratio from multiple 2X2 tables. J. Am. Statist. Assoc., 77, 145-152.

Hauck, W. W. and Wallemark (1983). A comparison of confidence interval methods for a common odds ratio. Presented at the August, 1993 Joint Statistical Meetings in Toronto, Ontario.

Jewell, N. P. (1984). Small-sample bias of point estimators of the odds ratio from matched sets, Biometrics, 40, 421-435.

Levin, B. and Kong, F. (1990). Bartlett's bias correction to the profile score function is a saddlepoint correction. Biometrika, 77, 219-221.

Lubin, J. H. (1981). An empirical evaluation of the use of conditional and unconditional likelihoods for case-control data, Biometrika, 68, 567-571.

Mantel, N. and Fleiss, J. L. (1980). Minimum expected cell size requirements for the Mantel-Haenzel one-degree-of-freedom chi-square test and a related rapid procedure, Amer. J. Epidemiol., 112, 129-134.

Mantel, N. and Haenzel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease, Journal of National Cancer Institute, 22, 719-748.

McKinlay, S. M. (1975). The effects of bias on estimators of relative risk for pair-matched and stratified samples, J. Am. Statist. Assoc., 70, 859-864.

McKinlay, S. M. (1978). The effect of nonzero second-order interaction on combined estimators of the odds ratio, Biometrika, 85, 191-202.

Paul, S. R. and Donner, A. (1989). A comparison of tests of homogeneity of odds ratios in K 2X2 tables, Statistics in Medicine, 8, 1455-1468.

Paul, S. R. and Donner, A. (1992). Small sample performance of tests of homogeneity of odds ratios in K 2X2 tables, Statistics in Medicine, 11, 159-165.

Robins, J. M., Breslow, N. E. and Greenland, S. (1986). Estimators of the Mantel-Haenzel variance consistent in both sparse data and large-strata limiting model, Biometrics, 42, 311-323.

Tarone, R. E. (1985). On heterogeneity tests based on efficient scores, Biometrika, 72, 91-95.

Tosia Sato (1990). Confidence limits for the common odds ratio based on the asymptotic ditribution of the Mantel-Haenzel estimator, Biometrics, 71-79.

Woolf, B. (1955). On estimating the relationship between blood group and desease, Annal. Human Genetics, 19, 251-253.

VITA AUCTORIS

The author was born in Jaffna Sri Lanka.

She received a B. Sc. degree in Mathematics and Physics from the University of Colombo, Sri Lanka in 1976.

She worked as a high school teacher in Sri Lanka and in Nigeria. (1976-1985)

She followed courses in Statistics at North Dakota State University. (1986)

She received a B. Ed. degree from the University of Windsor in 1994.