International Journal of Informatics and Data Science Research
ISSN 2997-3961 (Online)
Vol. 1, No. 10, Nov 2024,
Available online at: https://scientificbulletin.com/index.php/IJIDSR

Fisher Information Contained in Some Sets of Ordinal Statistics

Adirov Tolliboy Xasanovich
Professor of Tashkent State University of Economics, Tashkent, Uzbekistan

Article information:
Manuscript received: 4 Sep 2024; **Accepted:** 10 Oct 2024; **Published:** 14 Nov 2024

Abstract: Fisher information, a key concept in statistical theory, provides a measure of the amount of information that an observable random variable carries about an unknown parameter. In the context of ordinal statistics, Fisher information plays a crucial role in understanding the efficiency of parameter estimation. Ordinal data, which is categorized into ordered levels or ranks, presents unique challenges when it comes to statistical analysis, particularly in estimating parameters and evaluating the quality of statistical inference. This article explores Fisher information in the context of ordinal statistics, including methods for calculating Fisher information, results from recent studies, and an overview of relevant literature.

Keywords: Fisher information, ordinal data, ordinal statistics, including methods, results, overview.

1. Introduction

The concept of Fisher information arises from the field of statistical inference and plays a fundamental role in assessing the efficiency of estimators. It provides a quantitative measure of how much information a sample of data provides about an unknown parameter, with higher Fisher information corresponding to more precise estimates of the parameter. The formal definition of Fisher information, for a parameter θ, is given by the expected value of the negative second derivative of the log-likelihood function, which measures the curvature of the likelihood function and its sensitivity to changes in the parameter.

While Fisher information is well-established for continuous data, its application to ordinal statistics—where data consists of ordered categories rather than continuous values—requires additional considerations. Ordinal data presents specific challenges, such as the lack of an inherent distance between categories and the fact that the data may not fully follow standard distributional assumptions (e.g., normality). As a result, the calculation of Fisher information for ordinal data often requires adaptations of traditional methods or the development of new models tailored to the ordinal nature of the data.

In this article, we explore the methods used to calculate Fisher information for ordinal statistics, review key results from the literature, and highlight the challenges and opportunities in this area of research.

2. Ordinal Statistics: An Overview

Ordinal statistics deals with data that consists of categories with a natural ordering but without consistent spacing between categories. For instance, in a Likert scale survey, responses such as "strongly agree," "agree," "neutral," "disagree," and "strongly disagree" form ordered categories, but the exact distance between these categories is not defined. Despite this lack of cardinality, ordinal data is prevalent in many fields such as economics, psychology, social sciences, and medical research.

Statistical models for ordinal data are designed to account for the ordered nature of the variables. These models can be broadly divided into two categories:

Non-parametric models: These models do not make strong assumptions about the underlying distributions of the data and focus on the ranks or orders of the observations.

Parametric models: These models assume some form of underlying distribution for the ordinal data (e.g., multinomial or logistic distributions) and estimate parameters based on these assumptions.

Fisher information, typically applied in parametric models, can be used to assess the precision of parameter estimates derived from these models. Fisher information quantifies the sensitivity of the likelihood function to changes in model parameters, with larger Fisher information indicating that the parameter is estimated more precisely.

3. Fisher Information in Ordinal Statistics

Fisher information in the context of ordinal statistics is an extension of its classical definition, tailored to the properties of ordinal data. For a set of ordinal observations $Y_1, Y_2, \ldots, Y_n$, the Fisher information matrix I(θ) is defined as the expected value of the negative second derivative of the log-likelihood function:

$$\mathrm{I}(\theta) = -\,\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\log L(\theta, Y_1, Y_2, \ldots, Y_n)\right]$$

$L(\theta, Y_1, Y_2, \ldots, Y_n)$ is the likelihood function for the parameter θ\thetaθ based on the observed ordinal data.

In general, Fisher information in ordinal statistics can be more complex than in the case of continuous data due to the discrete nature of the data. Ordinal data often involves cumulative distribution functions (CDFs) or latent variable models that represent the ordering of categories. As a result, Fisher information for ordinal data is often computed using approximations or numerical methods. Some key challenges in calculating Fisher information for ordinal data include:

Lack of a distance metric: Ordinal data does not have a well-defined distance between categories, making it difficult to apply traditional Fisher information calculations.

Distribution assumptions: Models for ordinal data often rely on assumptions about the underlying distribution (e.g., cumulative logit models, proportional odds models), which can affect the calculation of Fisher information.

Complexity of likelihood functions: The likelihood functions for ordinal models may not have closed-form derivatives, requiring numerical methods for calculating Fisher information.

4. Methods for Calculating Fisher Information for Ordinal Data

Several methods have been developed to calculate Fisher information for ordinal data, each depending on the specific modeling assumptions and the nature of the data. Below, we highlight some common approaches used in the literature:

4.1 Cumulative Logit Models

One of the most widely used parametric models for ordinal data is the cumulative logit model, which assumes that the probability of an ordinal outcome is governed by a set of latent continuous variables. In a typical cumulative logit model, the probability of observing a response Y=k (where k is an ordinal category) is modeled as:

P(Y=k)= Φ $(\theta_k - \eta)$

Where $\boldsymbol{\Phi}$ is the cumulative distribution function of the standard normal distribution, θ_k are threshold parameters that define the cut-off points between categories, and η is a linear predictor based on covariates.

The Fisher information for the parameters in a cumulative logit model is calculated by taking the second derivative of the log-likelihood function with respect to the parameters and then computing the expected value. In practice, numerical methods are often used for this step, as the log-likelihood function is not always analytically tractable.

4.2 Proportional Odds Models

Another common model used for ordinal data is the proportional odds model, which assumes that the odds of being in a higher category (or below a threshold) follow a logistic distribution. The Fisher information for the parameters in a proportional odds model is similarly computed by taking the second derivative of the log-likelihood function with respect to the model parameters. This can be done numerically, especially when the log-likelihood function does not have a simple closed form.

4.3 Latent Variable Models

Latent variable models assume that ordinal responses arise from an underlying continuous latent variable, which is categorized into ordinal classes. For example, in a simple linear latent variable model, the observed ordinal variable *Y* is determined by the latent variable *X* through a set of cut-off thresholds:

$Y=k$ if $\theta_{k-1} < X \leq \theta_k$

In these models, Fisher information is derived by first calculating the likelihood of observing a specific ordinal category given the latent variable and then differentiating with respect to the model parameters.

4.4 Numerical and Monte Carlo Methods

In many cases, analytical expressions for Fisher information are unavailable or difficult to obtain for ordinal data models. As a result, numerical methods such as Monte Carlo simulations are often employed to estimate Fisher information. These methods involve simulating many datasets based on the model parameters and using numerical derivatives of the log-likelihood to approximate Fisher information.

5. Results from Recent Studies

Recent studies have examined the properties of Fisher information in ordinal statistics, focusing on its role in parameter estimation and model selection. Some notable findings include:

Efficiency of Ordinal Estimators: Studies have shown that models such as the cumulative logit model and proportional odds model can provide efficient estimators for ordinal data, with Fisher information providing a useful tool for assessing the precision of these estimates (Tutz & Schmid, 2016).

Impact of Sample Size: The Fisher information in ordinal models is often sensitive to the sample size. Smaller sample sizes typically result in larger variance in parameter estimates, as Fisher information is inversely related to the variance of the estimator (Gao et al., 2018).

Comparison of Ordinal and Continuous Models: Fisher information has been used to compare ordinal models with continuous data models, demonstrating that ordinal data models may sometimes require more complex estimation procedures due to the discrete nature of the data (Greene, 2017).

Use in Model Selection: Fisher information has been used in model selection criteria, such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), to compare different ordinal models and select the most appropriate model for a given dataset (Burnham & Anderson, 2004).

6. Conclusion

Fisher information plays an important role in the statistical analysis of ordinal data by quantifying the precision of parameter estimates. However, applying Fisher information to ordinal statistics is more complex than in continuous data due to the discrete nature of ordinal categories and the challenges in model specification. Methods such as cumulative logit models, proportional odds models, and latent variable models provide useful frameworks for calculating Fisher information in ordinal statistics. Numerical and simulation-based methods are often necessary for models where closed-form solutions are not available.

The ongoing development of Fisher information in ordinal data models is an active area of research, with applications in fields ranging from economics to biostatistics. As methods continue to evolve, Fisher information will remain an essential tool for assessing the efficiency of statistical estimators and guiding model selection in ordinal data analysis.

Literature

1. TS Safarov, SX Turakulov, IS Nabiyeva, SS Nabiyeva // Эффективность медицинские информационные системы в диагностике // Theoretical & Applied Science, 301-305.
2. Sakiev T., Nabieva S. // Architecture of the medical information system. International Scientific Journal Theoretical & Applied Science. Section 4. Computer science, computer injineering and automation. Issue: 05 Volume: 61. Published: 14/05/2018. p. 35-39
3. H.A. Primova, T.R. Sakiyev and S.S. Nabiyeva // Development of medical information systems// Journal of Physics: Conference Series. 1441 (2020) 012160 IOP Publishing doi:10.1088/1742- 6596/1441/1/012160 (Scopus) https://iopscience.iop.org/article/10.1088/1742-6596/1441/1/012160
4. Karshiev A., Nabieva S., Nabiyeva I. // Medical information systems. International Scientific Journal Theoretical & Applied Science. SECTION 4. Computer science, computer injineering and automation. Issue: 04 Volume: 72. Published: 30/04/2019. 505-508 p.
5. Mukhamedieva D.T., Primova Kh.A. // Approach to problem solving multicriterial optimization with fuzzy aim // International Journal of Mathematics and Computer Applications Research (IJMCAR) ISSN(P): 2249-6955; ISSN(E): 2249- 8060 Vol. 4, Issue 2, USA. 2014, 55-68 pp. Impact Factor (JCC): 4.2949
6. AB Karshiev, XA Primova, SS Nabiyeva, AS Egamkulov // Architectural integration problems of MIS // ISJ Theoretical & Applied Science, 05 (85), 733-739 p