**Bias Reduction of Maximum Likelihood Estimates STOR**

**Summary:**

The document discusses bias reduction in maximum likelihood estimation in regular parametric problems and logistic regression models. It introduces the concept of modifying the score function to reduce bias. The section also provides examples and illustrations of bias reduction, compares different adjustment procedures, and extends the concept to other generalized linear models. It highlights the use of the Jeffreys prior as a bias-reducing penalty function and mentions the need for further study on standard errors and confidence regions based on bias-reduced estimates.

**Bias reduction methods:**

This section of the document discusses bias reduction of maximum likelihood estimates in regular parametric problems. It introduces the concept of modifying the score function to reduce bias in the estimates. In the case of exponential families with canonical parameterization, the modification penalizes the likelihood by the Jeffreys invariant prior. The section also provides examples of bias reduction in the normal distribution and binomial logistic regression.

**Bias reduction methods:**

This section discusses bias reduction of maximum likelihood estimates in the context of logistic regression and other generalized linear models. It explains that the maximum penalized likelihood estimate is unique and exists due to the concave nature of the log-likelihood function.

The section also provides an illustration of shrinkage in a logistic regression example with a small sample size. It shows the bias-reduced estimates and the sampling distribution for different values of the true parameter.

Additionally, the section discusses the adjustment procedure for calculation of bias-reduced estimates and compares it to other adjustment procedures. It also extends the concept of bias reduction to other generalized linear models.

Finally, the section presents several examples, including the normal distribution, reciprocal mean of a Poisson distribution, normal distribution with known coefficient of variation, and precision of duplicate measurements. It demonstrates how the bias-reduced estimates can be calculated in each of these examples and discusses their properties.

**Bias reduction techniques:**

This section discusses bias reduction in maximum likelihood estimation. The Jeffreys prior is used as a bias-reducing penalty function in exponential family problems. The choice to reduce bias depends on the skewness of the maximum likelihood estimator and any sacrifice in precision. The section also mentions the use of observed and expected information for bias correction in parameterizations of exponential family models. The need for further study on standard errors and confidence regions based on bias-reduced estimates and the range of problems in which the bias-reduced estimator is guaranteed to be finite are mentioned as well. The Jeffreys prior is highlighted as having a special role in bias reduction.

**Tutorial on maximum likelihood estimation**

**Summary:**

The document provides a tutorial on maximum likelihood estimation (MLE) in statistics. It explains the concept of MLE, the likelihood function, and how it is used to estimate parameter values. It discusses the difference between MLE and least-squares estimation (LSE), and the optimal properties of MLE. The document also discusses non-linear optimization algorithms and the issue of local maxima. It presents an example of MLE applied to forgetting data. Finally, it compares MLE and LSE using a data set from a recall experiment and emphasizes the importance of model selection.

**Maximum likelihood estimation (MLE) tutorial**: explaining non-linear parameter estimation using non-normal data with examples and comparing to least-squares estimation (LSE)

This section of the document is a tutorial on maximum likelihood estimation (MLE). MLE is a method of parameter estimation in statistics that is used in non-linear modeling with non-normal data. The tutorial provides a conceptual explanation of MLE with illustrative examples. It explains the difference between MLE and least-squares estimation (LSE), stating that MLE has many optimal properties in estimation and is a standard approach in statistics, while LSE is primarily used with linear regression models and has no basis for testing hypotheses or constructing confidence intervals. The tutorial also introduces the concept of probability density function (PDF) and likelihood function, and explains how they are used in model specification and parameter estimation.

**MLE explained in detail:**

This section discusses the concept of maximum likelihood estimation (MLE). It explains that MLE is a method used to find the parameter values that make the observed data most likely. The likelihood function, which represents the probability distribution that underlies the data, is maximized to find the MLE estimate. The section also discusses the likelihood equation and how it represents a necessary condition for the existence of an MLE estimate. It explains that non-linear optimization algorithms are typically used to find the MLE estimate when an analytic solution is not possible. The section also mentions the issue of local maxima, where the optimization algorithm may find sub-optimal solutions. It discusses how different starting parameter values can lead to different solutions and suggests techniques to avoid the local maxima problem. Additionally, the section discusses the relationship between MLE and least-squares estimation (LSE). It explains that while MLE seeks the parameter values that make the data most likely, LSE seeks the parameter values that minimize the sum of squares error between observations and predictions. The section concludes with an illustrative example of MLE applied to forgetting data.

**MLE vs LSE analysis:**

This section of the document discusses the process of maximum likelihood estimation (MLE) and compares it to least-squares estimation (LSE). The section also presents the results of applying MLE and LSE to a data set from an experiment that involved subjects recalling words or letters. The results show that the exponential model fits the data better than the power model. The section concludes by discussing the importance of model selection based on generalizability.

**Maximum Likelihood Estimation of Logistic Regression Models- Theory and Implementation**

**Summary:**

The document provides a comprehensive overview of logistic regression models and the maximum likelihood estimation method. It discusses both binomial and multinomial logistic regression models, explaining the likelihood function and the parameter estimation process for each. The Newton-Raphson method is introduced as a way to numerically estimate the parameters, and cautionary notes are given regarding parameter estimates and convergence. The document concludes by providing a skeletal implementation of logistic regression using the C programming language.

**Logistic regression model:**

This section of the document introduces logistic regression as a model for dependent variables with discrete categorical levels. It explains that logistic regression is used when the response values are not measured on a ratio scale and the error terms are not normally distributed. The section also outlines the maximum likelihood estimation method for logistic regression models, which involves finding the parameters that best fit the data. Specifically, it focuses on binomial logistic regression for a binary dependent variable and discusses the model, parameter estimation, and the likelihood function. The section concludes by mentioning a generic implementation of the algorithm to estimate logistic regression models.

**Maximum likelihood estimation:**

This section of the document discusses the maximum likelihood estimation of logistic regression models. It explains that each solution specifies a critical point that is either a maximum or a minimum. The matrix of second partial derivatives determines if the critical point is a maximum or not. The matrix also forms the variance-covariance matrix of the parameter estimates. The Newton-Raphson method is described as a way to numerically estimate the parameter estimates for logistic regression models. It mentions cautionary notes about parameter estimates tending to infinity and the possibility of the iteration process not converging. Finally, it introduces multinomial logistic regression and explains how to compute the probabilities of the different categories of the dependent variable.

**Logistic regression estimation:**

This section of the document discusses the maximum likelihood estimation of logistic regression models. It explains the parameter estimation process for multinomial logistic regression models and provides equations for calculating the likelihood function and its first and second derivatives. The Newton-Raphson method is then used to iteratively maximize the likelihood function and estimate the parameters. The section also includes an outline of a skeletal implementation of logistic regression using the C programming language.