Frequentist inference

(6) Foundations and maximum likelihood

Study questions

1. Define the notion of a point estimator.
2. Write down the definitions of the likelihood and log likelihood functions.
3. Write down the definition of the maximum likelihood estimator.
4. Write down the general maximum likelihood procedure for parametric statistical models.
5. Derive the maximum likelihood estimator for the parameter of a Bernoulli distribution.
6. Derive the maximum likelihood estimator for the parameter of Gaussian distribution.
7. Formulate the univariate Newton-Raphson method
8. Formulate the multivariate Newton-Raphson method
9. Write down the definition of the score function.
10. Write down the definition of the Fisher information matrix
11. Write down the definition of the expected Fisher information.
12. Write down the Fisher scoring algorithm.

Exercises (Theory)

1. Derive the maximum likelihood estimator for the continuous uniform distribution on an interval $[0, \theta]$ (DeGroot and Schervish (2012, Examples 7.5.7), Wasserman (2004, Example 9.12)).
2. Develop the methods of moments for parameter estimation (e.g., DeGroot and Schervish, 2012, Definition 7.6.3, Examples 7.6.8. 7.6.9).
3. Evaluate the Fisher information and the expected Fisher information for the expectation parameter of a univariate and a multivariate Gaussian random variable, respectively (DeGroot and Schervish, 2012, Examples 8.8.3, 8.8.13, and 8.8.14).

Exercises (Programming)

1. Write a program that implements a Fisher scoring algorithm for the maximum likelihood estimation of the parameters of a simple linear regression model. Compare the results with the analytical estimation of the parameters. Visualize and document your results.
2. Let $X_1, ..., X_n \sim \text{Bern}(\mu)$ be $n = 20$ i.i.d. Bernoulli random variables. Using an optimization routine of your choice, formulate and implement the numerical maximum likelihood estimation of μ for true, but unknown values of $\mu = 0.7$ and $\mu = 1$ based on $X_1, ..., X_n$.
3. Let $X_1, ..., X_n \sim \text{Bern}(\mu)$. For a large number n, sample the $X_1, ..., X_n$ and evaluate the maximum likelihood estimator $\hat{\mu}^{ML}$. Repeat this m times and create a histogram of the realized $\hat{\mu}_1^{ML}, ..., \hat{\mu}_m^{ML}$. Visualize and document your results.