TMA4295 Statistical inference

Lecturer Fall 2023: *Gunnar Taraldsen*

Norwegian University of Science and Technology
Department of Mathematical Sciences

Scribe: *Johan Vik Mathisen* and *Maiken Hatten*

Lecture 1 in week 38: Estimation

Contents

1 Introduction and discussion or other titles

2 Fundamental concepts

Definition 2.1 (Estimator) An estimator is a statistic.

Definition 2.2 (Predictor) A predictor is a statistic.

An estimator is a random point; an estimate is a realized value. Similarly, a predictor is a random point and a prediction is a realized value. Intuition: an estimate (and also a prediction) is a guess for something unknown.

We have two cases: (1) an estimate for a parameter, and (2) a prediction of a random point X, where we have orthogonality such that

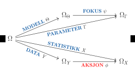
$$\phi(W) \perp (X - E(X|W)) \tag{1}$$

Definition 2.3 (?) $E(X\phi(W)) = E(E(X|W)\phi(W))$ defines $E(X(W))$ uniquely.

Example 2.1 (Prediction of the future) Given a graph consisting of observations W and unknowns X, P_X^w is the conditional distribution of X given P. Furthermore,

$$P_X^w(A) = P(X \in A | W = w) \tag{2}$$

such that P_X^w is the state of knowledge about X given W.



2.1 Median

We can use the median of X, denoted $M^w(X)$, as a predictor of X instead of $E^w(X)$. $\mu = E^w(X)$ minimizes $E^w(X-\mu)^2$, and $\mu_m = M^w(X)$ minimizes $E^w|X-\mu_m|$.

Note:

$$E(f(X)) \neq f(E(X)) \tag{3}$$

If f is 1-1, then

$$M(f(X)) = f(M(X)) \tag{4}$$

3 Estimation

1 **Types of estimates**

a) Point estimate $t = t(x) \in \Omega_T$, where Ω_T is the target space of the parameter $\tau = \tau(\theta)$

b) Set estimate

c) Distribution estimate

2 **Methods for construction of point estimates**

a) Maximum likelihood estimation (MLE)

b) Bayes

c) Empirical distribution, P_n (non-parametric bootstrap estimation)

d) Moment estimates

e) Rao-Blackwell

Think of point estimates as measurement instruments.

Example 3.1 (Maximum likelihood) Let $X_1, \ldots, X_n \sim \text{Exp}(\beta)$. The likelihood L is defined as

$$L = \prod_{i=1}^{n} \frac{1}{\beta} e^{\frac{-X_i}{\beta}} \tag{5}$$

$$= \beta^{-n} \cdot e^{-\bar{X} \cdot \frac{n}{\beta}} \tag{6}$$

$$= \left(e^{\frac{-\bar{X}}{\beta}} \cdot \frac{\bar{X}}{\beta} \right)^n \cdot \bar{X}^{-n} \tag{7}$$

MLE is achieved when $\frac{\bar{X}}{\beta} = 1$. Thus $\hat{\beta} = \bar{X}$ is MLE.

Example 3.2 (Bayes) (There are probably mistakes here)
Let $X_i \sim \text{Exp}(\beta) = \Gamma(\alpha = 1, \beta)$ and $\bar{X} \sim \Gamma(\alpha = n, \beta/n)$.

We choose $\pi(\beta) = 1/\beta$ as improper prior. The posterior is given by $\pi(\beta|x) \propto L(\beta)\pi(\beta)$ whose functional part is given by

$$L(\beta)\pi(\beta) = \beta^{-(n+1)} \cdot e^{-\bar{X}n\beta^{-1}} \tag{8}$$

As a function of β this is the functional part of the inverse gamma pdf with parameters: shape $\alpha = n$ and scale $\beta = n\bar{X}$.

An estimator of the value of the parameter of the original exponential model is the expected value of β.

$$\beta^\star = \text{E}(\beta) = \frac{n\bar{X}}{n-1} = c\bar{X}$$

Example 3.3 (Rao-Blackwell) Idea: Take a crude estimator and project to make it better!

Let $\bar{X}$ be sufficient (and unbiased?). Rao-Blackwell estimator for $\bar{X} = \text{E}(\bar{X}|\bar{X}) = \bar{X}$.

For the sample variance S^2 we have that $\text{E}(S^2|\bar{X}) \neq S^2$, but is now optimal in the MSE sense, although his is not necessarily a natural demand.

Definitions

Theorems

Examples

Index