Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4295 Statistical inference

Lecturer Fall 2023: *Gunnar Taraldsen*

Scribe: *Johan Nikolai Kristiansen*

Lecture 20 in week 43: 'Set estimation'

Contents

1 Repetition

Example 1.1 (Squared logarithmic error loss) Assume a statistic $t = t(x)$ where x is the data. The easiest method of estimation will often be point estimation ($x \in \mathbb{R}$), with loss function $(l = \ln(t) - \ln(\tau))^2$ where l would be squared logarithmic error loss. The difference between this squared error loss and normal, is that this includes the logarithm which measures scale invariant apposed to invariance shift.

Example 1.2 (Loss related to type 1 and type 2 error) A test only gives two values, 0 and 1 (recall definition of a test).
If $t = 1 \Rightarrow H_1$ true, and $t = 0 \Rightarrow H_0 true$
If $\tau(1-t)$, $\tau = 1$ and $t = 1 \Rightarrow 0$ loss, $\tau = 1$ and $t = 0 \Rightarrow$ loss. This gives the general equation

$$\tau(1-t) + \lambda(1-\tau)t \tag{1}$$

where $\tau(1-t)$ corresponds to type 2 error and $\lambda(1-\tau)t$ corresponds to type 1 error.

Example 1.3 (Bayes posterior distribution estimator)

$$t(d\tau) = \pi(\tau|x)d(\tau) \tag{2}$$

2 Confidence sets

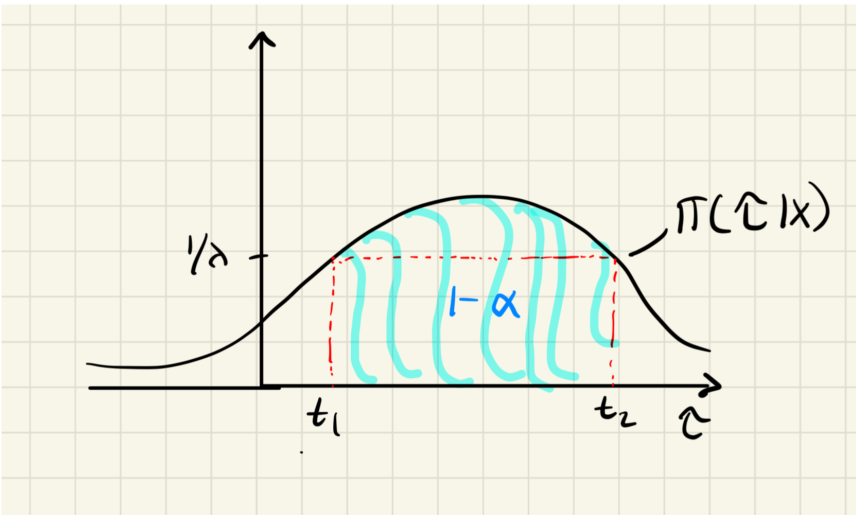

Figure 1: The shortest credibility interval with level = 1 - α

Here we chose an interval (red dots) and the density is equal at t_1 and t_2 if $\pi(|x)$ is **unimodal**

Definition 2.1 (Unimodal) A function is unimodal if the right and side of the interval is non increasing and the left hand side is non decreasing. This gives the shortest interval (proven in the book, and below).

Definition 2.2 (Confidence level) A set estimator S has confidence level $1 - \alpha$ if $P(\tau) \geq 1 - \alpha$, where P is the hitting probability, τ is the parameter and S is a random set.

Theorem 2.1 () Optimal Bayes credibility intervals from the right Haar prior are frequentist optimal (Equivariant!)

Proof. Firstly we need to define some mathematical properties. Let $x = \theta * u \Leftrightarrow \theta = xu^{-1}$ which gives $x = \Theta^x * U$ where $\Theta^x = x * U^{-1}$, posterior.
We also have the invariance: $t(x) = t(\theta * u) = \theta * t(u) \Rightarrow l(\theta * t, \theta * \tau) = l(t, \tau)$.
We need $|\theta * t| = |t| =$ size of t

Then we can start formulating the proof.
The risk can be written as

$$\rho = El(T, \tau) = El(t(\theta * U), \tau(\theta)) = El(t(x), \tau(\Theta^x)) = \rho^x \tag{3}$$

which equals Bayes posterior risk from this right invariant prior, where $\Theta^x \sim$ posterior of θ given data x ☐

Example 2.1 $x_1, ..., x_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\theta, \sigma^2)$ where $\theta + z * \sigma_i$. Can also be formulated as $\bar{x} = \tau + \sigma * \bar{z}, a_i = x_i - \bar{x}$.
We need to condition $\bar{z}$ by a_i: $\bar{z}|a \sim \bar{z}$ from Basu.

Example 2.2 (More general) $x = \theta - u$ where u is a known distribution
$\Theta^x = x + U \sim$ Bayes posterior. Then we need to calculate the loss and the risk.
$l = |s| + \lambda(\theta \notin s)$
$\rho^x = |s| + \lambda * P(x + U \notin s)$
$s(x) = x + s(0)$ by assumtion, $s(0) = [t_1, t_2]$
$t_2 - t_1 + \lambda * P(U \notin [t_1, t_2] = t_2 - t_1 + \lambda * [F(t_1) + (1 - F(t_2))]$
Minimize: $U \sim f = F'$ gives us two equations with two unknowns:
$0 = \partial_1[t_2 - t_1 + \lambda * [F(t_1) + (1 - F(t_2))]] = -1 + \lambda * f(t_1)$
$0 = \partial_2[t_2 - t_1 + \lambda * [F(t_1) + (1 - F(t_2))]] = 1 - \lambda * f(t_2)$
$0 = \lambda * (f(t_1) - f(t_2) \Rightarrow f(t_1) = f(t_2) = 1/\lambda$
The level is $1 - \alpha$ where
$\alpha = P(\theta \notin S) = P(\theta \notin (X + [t_1, t_2])) = P(U \notin [t_1, t_2])$
Note: Minimizing ρ^x gives also the minimal size $|s|$ under the constraint
$P_x(\Theta^x \in s) \geq 1 - \alpha$

To summarize the methods:

(1) Invert likelihood test

(2) Optimal equivariant from Bayes posterior risk

(3) Pivotal quantity ($x = \theta - u \Leftrightarrow x - \theta = -u$ is pivotal)

Definition 2.3 (Pivotal point) A pivotal point $Q = Q(X, \theta)$ has a known distribution, where X is the data and θ is the model

Example 2.3 $(\bar{x} - \mu)/s$ is a pivot in $x_i = \mu + \sigma * z_i \Rightarrow \theta = (\mu, \sigma)$

Example 2.4 (Standard example) $\bar{x} \pm ku$ is a $1 - \alpha$ confidence interval for μ given, and $x_i \sim N(\mu, \sigma^2)$ where $k = t_{\alpha/2, n-1}, \sigma$ unknown.
$k = z_{\alpha/2}, \sigma$ known
ku = expanded uncertainty
u = $\begin{cases} \frac{s}{\sqrt{n}} & \sigma \text{ is unknown} \\ \frac{\sigma}{\sqrt{n}} & \sigma \text{ is known} \end{cases}$ where u = standard uncertainty and k = coverage factor
$\left(\frac{\bar{x}-\mu}{u}\right) \sim \begin{cases} t_{(n-1)} & \text{when } \sigma \text{ is unknown} \\ \mathcal{N}(0,1) & \text{when } \sigma \text{ is known} \end{cases}$

Definitions

Theorems

Examples

Index