# TMA4295 Statistical inference

Lecturer Fall 2023: *Gunnar Taraldsen*

Scribe: *Ola Rasmussen*

Norwegian University of Science and Technology
Department of Mathematical Sciences

**Lecture 1 in week 40: 'Standard Uncertainty'**

## Contents

# 1 Fundamental Concepts

**Definition 1.1 (Statistic)** A statistic is a function of data.

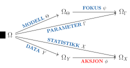
**Definition 1.2 (Parameter)** A parameter is a function of model.

# 2 Standard Uncertainty

**Definition 2.1 (Standard Uncertainty)** Standard uncertainty $u$ is an estimate of the standard deviation of the estimator.

One should always state the Standard Uncertainty!

**Example 2.1 (Uncertainty of the length of the NTNU pendulum)** The length of the NTNU pendulum is $25.26(1)m$, where $25.26m$ is the estimate of the length, and (1) is the standard uncertainty. In this context, $25.26(1)m = 25.26m \pm 0.01m$.




**Example 2.2 (Estimators of the length of the NTNU pendulum)** The length of a pendulum is given by,

$$\lambda = \left(\frac{\tau}{2\pi}\right)^2 g, \tag{1}$$

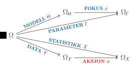
where $\tau$ is the period and $g$ is the earths gravitational pull. Two possible estimates of this parameter are;

$$\hat{\lambda} = \left(\frac{\bar{t}}{2\pi}\right)^2 g, \tag{2}$$

$$\lambda^* = \bar{x}, \text{ where } x_i = \left(\frac{t_i}{2\pi}\right)^2 g, \tag{3}$$

These two estimators are two different instruments used for estimating the length. So the question is, which one do we pick?

To decide this, we need a specific statistical model. One possibility could be that $t_1, \ldots, t_n \sim \mathcal{N}(\mu, \sigma^2)$. Then the best estimator would be the one that is minimally sufficient. In this possibility, the common minimally sufficient statistics are $\bar{t}$ and $s$. Since $\hat{\lambda}$ is the only estimator that depends on $\bar{t}$, it is the best one of the two. $\lambda^*$ does not, but maybe this estimator could give a smaller bias or variance than $\hat{\lambda}$. The Rao-Blackwellization, $\lambda^{**} = E[\lambda^*(T)|\bar{t}, s_{(t)}]$, of $\lambda^*$ could give us a third estimator, and this estimator would be better than $\lambda^*$.




How do we find the optimal estimator of $\lambda$? The distribution of $\lambda$ is found through,

$$\bar{t} = \tau + \sigma\bar{z}, \text{ where } z_i \sim \mathcal{N}(0, 1), \tag{4}$$

$$\bar{t}^2 = \tau^2 + 2\sigma\tau\bar{z} + \sigma^2\bar{z}^2, \tag{5}$$

$$E[\bar{T}^2] = \tau^2 + \sigma^2 E[\bar{z}^2], \tag{6}$$

$$E[\bar{T}^2 - ks^2] = \tau^2. \tag{7}$$

This gives UMVU (uniformly minimum-variance unbiased) and UMRU (uniformly minimum-risk unbiased) estimators of $\lambda$:

$$\check{\lambda} = \frac{g}{(2\pi)^2}(\bar{t}^2 - ks^2). \tag{8}$$

# 3 Lower bound of the standard deviation

The Cramer-Rao inequality gives us,

$$Var[T] \geq (\tau')^2 \iota^{-1}. \text{ where } \iota = Var[S] = Var[\partial_\theta ln(f(X))]. \tag{9}$$

So $\sqrt{Var[T]} \geq \sqrt{\tau^2 \iota^{-1}}$ gives the lower bound of the standard deviation of $T$, and $T$ is unbiased. This is easier to calculate than the standard deviation. One idea could be to use the Cramer-Rao inequality to report the standard uncertainty, but this would be optimistic. The idael case would be that $\tau(\theta) = a + b\theta$, then $\tau(\hat{\theta}) = a + b\hat{\theta}$. Then,

$$Var[\tau(\hat{\theta}(X))] = b^2 Var[\hat{\theta}(X)]. \tag{10}$$

Here, $\tau(\hat{\theta}(X))$ is only unbiased if $\hat{\theta} \sim \mathcal{N}(0, \sigma)$, and then $Var[\hat{\theta}(X)] = \iota^{-1}$

# 4 Fisher Information Metric

Let $X \sim f$, $R(X) = \{x_1, \dots, x_m\}$, $X$ is a simple random point. $f(x) = P(X = x)$, and $f$ is unknown, $\sqrt{f} \in \Re^m$, $|\sqrt{f}| = 1$.

Assume $P_X$ is known when time $t$ is known. When $m = 3$ we have the unit ball. Distance is given by speed and time, $d(t_1, t_2) = \int_{t_1}^{t_2} v dt$. Then,

$$\begin{aligned}
v^2 &= \sum_{i=1}^{m} (\partial_t \sqrt{f(x_i)})^2, \text{ where } \partial_t \sqrt{f} = \frac{1}{2} \partial_t ln(f) \cdot f^{\frac{1}{2}} \\
&= \sum_{i=1}^{m} \frac{1}{4} (\partial_t ln(f))^2 f \\
&= \frac{1}{4} E[S^2], \text{ where } S \text{ is the score function.}
\end{aligned}$$

$$v = \frac{1}{2} \sqrt{\iota}, \text{ where } \iota = E[S^2] = Var[S].$$

So, $d(t_1, t_2) = \frac{1}{2} \int_{t_1}^{t_2} \sqrt{\iota} dt$. In conclusion, the Fisher Information gives both a method for calculating Standard Uncertainty and also a metric on the model space.

# Definitions



# Theorems

# Examples

# Index