Ausgewählte Kapitel der Parameterschätzung — Übung 4

Ausgabedatum: 08. June 2021 Abgabetermin: 22. June 2021, 18 Uhr

Wahrscheinlichkeit und der Satz von Bayes

A gambler in a casino is surprised by the outcomes of a repeated coin-tossing and would like to know if the coin is fair, that is, 50% of chance to observe a head and 50% to observe a tail. The gambler proceeds to a simple experiment by tossing the coin 200 times. The results are reported in the file `cointoss.mat` where 0 stands for tail and 1 for head.

a) Following the frequentist approach, suggest a simple method to determine whether the coin is fair or not.

b) To ascertain and quantify this conclusion, we will also follow a Bayesian approach. Let us note $H \in [0, 1]$ the bias-weighting of the coin. $H = 0$ represents a coin which produces a tail at every flip, $H = 1$ a head at every flip and $H = 0.5$ represents a fair coin. Intuitively, H can be interpreted as the probability to obtain a head by tossing this coin. We would like now to estimate the posterior pdf (probability density fonction) $prob(H|\{results\})$. Use Bayes' theorem to relate it to two other pdf, which are easier to assign. Recall which one is the likelihood function and which one is the prior pdf.

c) Suppose *results* is the proposition "R heads in N tosses". What is the probability $prob(\{results\}|H)$?

d) Since we do not have any a priori about whether the coin is fair or not, we will suppose that $prob(H) = 1$ if $0 \leq H \leq 1$ and $prob(H) = 0$ otherwise. Compute and plot $prob(H|\{results\})$ for $N = 1, 2, 3$ and 4. Comment the evolution of the pdf.

e) Compute and plot $prob(H|\{results\})$ for $N = 200$ and conclude.

f) A possible question is the following: Can the posterior pdf computed for $N = k$ be used as prior for $N = k + 1$, that is sequentially? If yes, compute the final pdf.

g) Actually, the first question led us to think that the coin is not fair and rather biased toward heads. Therefore, we use another prior pdf, $prob(H) = 3H^2$ if $0 \leq H \leq 1$ and $prob(H) = 0$ otherwise. Follow question e) and compare your results with the ones obtain with a uniform prior.

Die Maximum-Likelihood-Methode

Maximum likelihood estimation is a common method of statistical inference to estimate the unknown parameters of a probability distribution. Suppose we observe a sample $(x_1, x_2, ..., x_n)$ of n random variables $\mathbf{X} = (X_1, X_2, ..., X_n)$. Let's call $f(\mathbf{X}|\theta)$ the joint pdf of X, which depends on a set of parameters $\theta = (\theta_1, ..., \theta_k) \subset \Theta$. The θs can be parameters such as mean or variance of some pdf. The **likelihood function** $\mathcal{L}_n$ for a given sample is the evaluation of the joint pdf at the observed sample:

$$\mathcal{L}_n(\theta) = f((x_1, x_2, ..., x_n)|\theta) \tag{1}$$

We define also the **log-likelihood function** as

$$l_n(\theta) = \log \mathcal{L}_n(\theta) \tag{2}$$

The **maximum likelihood estimator** (abb. MLE) $\hat{\theta}_n$ of θ is, when it exists and is unique

$$\hat{\theta}_n = \underset{\theta}{\operatorname{argmax}} \mathcal{L}_n(\theta). \tag{3}$$

Equivalently, it is also the parameter θ that maximize the log-likelihood function. In particular, if we assume that the X_i $i \in 1, \ldots, n$ are independent random variables of respective pdf $f_i(X_i, \theta)$ then

$$\mathcal{L}_n(\theta) = \prod_{i=1}^{n} f_i(x_i|\theta) \tag{4}$$

a) What is the maximum-likelihood estimation of H?

Karim Douch — Raum 5.306 — karim.douch@gis.uni-stuttgart.de