WuS - Lecture Notes Week 10

Ruben Schenk, ruben.schenk@inf.ethz.ch

June 24, 2022

0.1 Die Maximum-Likelihood-Methode (ML-Methode)

Ausgangspunkt im folgenden Abschnitt ist immer eine von zwei Situationen, jenachdem ob wir es mit diskreten oder mit stetigen Zufallsvariablen zu tun haben. Wir schreiben oft kurz $\vec{X} = (X_1, ..., X_n)$. In jedem Modell $\mathbb{P}_\theta$ sind $X_1, ..., X_n$ entweder diskret mit gemeinsamer Gewichtsfunktion $p_{\vec{X}}(x_1, ..., x_n; \theta)$ oder stetig mit gemeinsamer Dichtefunktion $f_{\vec{X}}(x_1, ..., x_n; \theta)$. Meistens sind sogar die X_i unter $\mathbb{P}_\theta$ i.i.d. mit individueller Gewichtsfunktion $p_X(x; \theta)$ bzw. Dichtefunktion $f_X(x; \theta)$. Dann ist also die gemeinsame Gewichtsfunktion

$$p_{\vec{X}}(x_1, ..., x_n; \theta) = \prod_{i=1}^{n} p_X(x_i; \theta)$$

bzw. die gemeinsame Dichtefunktion

$$f_{\vec{X}}(x_1, ..., x_n; \theta) = \prod_{i=1}^{n} f_X(x_i; \theta).$$

Anschaulich ist

$$p_{\vec{X}}(x_1, ..., x_n; \theta) = \mathbb{P}_\theta[X_1 = x_1, ..., X_n = x_n]$$

gerade die Wahrscheinlichkeit im Modell $\mathbb{P}_\theta$, dass unsere Strichprobe $X_1, ..., X_n$ die Werte $x_1, ..., x_n$ liefert, und $f_X(x_1, ..., x_n; \theta)$ ist das übliche stetige Analog.

Def: Die **Likelihood-Funktion** ist:

$$L(x_1, ..., x_n; \theta) := \begin{cases} p_{\vec{X}}(x_1, ..., x_n; \theta) & \text{im diskreten Fall,} \\ f_{\vec{X}}(x_1, ..., x_n; \theta) & \text{im stetigen Fall.} \end{cases}$$

Die Funktion $\log L(x_1, ..., x_n; \theta)$ heisst die **log-Likelihood-Funktion.** Sie hat gegenüber der Likelihood-Funktion den Vorteil, dass sie im i.i.d.-Fall durch eine Summe (statt ein Produkt) gegeben und damit zum Rechnen oft wesentlich einfacher ist.

Def: Für jedes $x_1, ..., x_n$ sei $t_{ML}(x_1, ..., x_n) \in \mathbb{R}$ der Wert, der $\theta \to L(x_1, ..., x_n; \theta)$ als Funktion von θ maximiert. D.h.,

$$L(x_1, ..., x_n; t_{ML}(x_1, ..., x_n)) = \max_{\theta \in \Theta} L(x_1, ..., x_n; \theta).$$

Ein **Maximum-Likelihood-Schätzer (ML-Schätzer)** T_{ML} für θ wird definiert durch

$$T_{ML} = t_{ml}(X_1, ..., X_n).$$

Meistens sind $X_1, ..., X_n$ i.i.d. unter $\mathbb{P}_\theta$. Die Likelihood-Funktion L ist dann ein Produkt, und es ist bequemer, statt L die log-Likelihood-Funktion $\log L$ zu maximieren, weil diese eine Summe ist. Statt zu maximieren sucht man ferner meistens nur *Nullstellen der Ableitung (nach θ).*