Karlsruher Institut für Technologie

Institut für Volkswirtschaftslehre (ECON)

Lehrstuhl für Ökonometrie und Statistik

Kapitel VI - Maximum-Likelihood-Schätzfunktionen

Induktive Statistik

Prof. Dr. W.-D. Heller
Hartwig Senska
Carlo Siebenschuh

In komplexen Schätzsituationen ist eine GBE-Schätzfunktion nicht immer zu bestimmen.

Daher betrachten wir noch ein weiteres Schätzprinzip, das intuitiver ist.

Ausgangssituation:

Gegeben seien Beobachtungswerte $x_1, x_2, \ldots, x_n$ einer Stichprobe

(nicht notwendig mit Zurücklegen).

$X = (X_1, X_2, \ldots, X_n)$ sei der zugehörige Zufallsvektor

X **diskret:**
Die Wahrscheinlichkeitsverteilung von X hängt vom zu schätzenden Parameter ab:

$P_\gamma(X = x)$	Wahrscheinlichkeit für die Beobachtung $x = (x_1, x_2, \ldots, x_n)$ beim Parameter γ

Idee: Wähle den Parameterwert als Schätzwert, bei dem die Beobachtung am plausibelsten ist.

$\hookrightarrow$ Maximum-Likelihood-Prinzip:

Maximiere $P_\gamma(X = x)$ *über* $\gamma \in \Gamma$
Schätzwert $\hat{\gamma}$=*Maximalstelle* γ^*

Beispiel: Poissonverteilung
Bei einer einfachen Stichprobe mit Zurücklegen ist

$$P_\lambda(X = x) = \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} = \lambda^{\sum\limits_{i=1}^{n} x_i} e^{-n\lambda} \frac{1}{x_1! x_2! \dots x_n!}$$

zu maximieren über λ. $\Rightarrow$ Notwendige Bedingung:

$$\frac{d}{d\lambda}\left(\lambda^{\sum\limits_{i=1}^{n} x_i} e^{-n\lambda}\right) = \sum_{i=1}^{n} x_i \lambda^{\sum\limits_{i=1}^{n} x_i - 1} e^{-n\lambda} - \lambda^{\sum\limits_{i=1}^{n} x_i} n e^{-n\lambda} = \lambda^{\sum\limits_{i=1}^{n} x_i} e^{-n\lambda} \left(\frac{\sum\limits_{i=1}^{n} x_i}{\lambda} - n \right) = 0$$

Daraus folgt für die Maximalstelle λ^* (sofern es eine gibt):

$$\lambda^* = \frac{1}{n} \sum_{i=1}^{n} x_i$$

(λ^* ist Maximalstelle, da die zweite Ableitung an dieser Stelle negativ ist.)

Maximum-Likelihood-Schätzwert: $\hat{\lambda}_{ML} = \frac{1}{n} \sum_{i=1}^{n} x_i$

X stetig:

Die Dichtefunktion von X hängt vom zu schätzenden Parameter ab:
$f_{X,\gamma}(x)$

> *Dichtefunktion von X an der Stelle $x = (x_1, x_2, \ldots, x_n)$ beim Parameter γ*

Idee:

> *Die Wahrscheinlichkeit für einen schmalen Bereich um x entspricht dem Wert der Dichtefunktion an der Stelle x, wird also maximal, wenn die Dichtefunktion ihren Maximalwert hat. Also wähle den Parameterwert γ als Schätzwert, bei dem die Dichtefunktion ihr Maximum hat.*

$\hookrightarrow$ **Maximum-Likelihood-Prinzip:**

> *Maximiere $f_{X,\gamma}(x)$ über $\gamma \in \Gamma$*
> Schätzwert $\hat{\gamma}$=Maximalstelle γ^*

Beispiel: Exponentialverteilung
Dichtefunktion einer einfachen Stichprobe mit Zurücklegen

$$f_{X,\lambda}(x_1, \ldots, x_n) = \begin{cases} 0 & x_i \leq 0 \text{ für ein } \ i \\ \\ \lambda^n \cdot e^{-\lambda \sum\limits_{i=1}^{n} x_i} & x_i > 0 \text{ für } \ i = 1, \ldots, n \end{cases}$$

Sei $x_i > 0$ für $i = 1, \ldots, n$.
Durch Logarithmieren der Dichtefunktion ändert sich die Maximalstelle nicht:

$$\ln f_{X,\lambda}(x) = n \cdot \ln \lambda + (-\lambda \sum_{i=1}^{n} x_i)$$

Damit ist

$$\frac{d}{d\lambda}(\ln f_{X,\lambda}(x)) = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i.$$

Die zweite Ableitung ist negativ, Maximalstelle ist also

$$\hat{\lambda}_{ML} = \frac{n}{\sum\limits_{i=1}^{n} x_i}$$

Maximum-Likelihood-Schätzwert ist also der Kehrwert des arithmetischen Mittels.
Dieses Schätzverfahren ist jedoch nicht erwartungstreu!

Allgemeine Vorgehensweise beim Maximum-Likelihood-Schätzverfahren

Sei X eine Stichprobe zur Zufallsvariable Y mit parametrischer Verteilungsannahme mit Parameterraum Γ und $\mathcal{X}$ der Stichprobenraum zu X.

1. Zu einem Stichprobenergebnis $x \in \mathcal{X}$ und $\gamma \in \Gamma$ sei

$$L_x(\gamma) \quad := \quad \begin{cases} P_\gamma(X = x) & \text{falls } Y \text{ und damit } X \text{ diskret} \\ f_{X,\gamma}(x) & \text{falls } Y \text{ und damit } X \text{ stetig} \end{cases}$$

$L_x(\gamma)$ heißt Likelihood-Funktion zu $x \in \mathcal{X}$.

Allgemeine Vorgehensweise beim Maximum-Likelihood-Schätzverfahren

2. Die Maximalstelle der Likelihood-Funktion zu $x \in \mathcal{X}$, d.h.

$$\hat{\gamma}_{ML} \text{ mit } L_x(\hat{\gamma}_{ML}) = \max_{\gamma \in \Gamma} L_x(\gamma)$$

heißt *Maximum-Likelihood-Schätzwert* (kurz: *ML-Schätzwert*) *zum Stichprobenergebnis* $x \in \mathcal{X}$.

Allgemeine Vorgehensweise beim Maximum-Likelihood-Schätzverfahren

Zur Bestimmung der Maximalstelle wird die Likelihood-Funktion häufig logarithmiert. Die notwendige Bedingung für die Maximalstelle lautet dann:

$$\frac{d}{d\gamma}(\ln(L_x(\gamma))) = 0$$

Diese Gleichung wird als ML-Gleichung bezeichnet. (Hat $\gamma = (\gamma_1, \ldots, \gamma_k)$ mehrere Komponenten (vgl. Normalverteilung und Gammaverteilung), so sind die partiellen Ableitungen gleich Null zu setzen und wir erhalten mehrere ML-Gleichungen.)

Allgemeine Vorgehensweise beim Maximum-Likelihood-Schätzverfahren

3. Die Schätzfunktion $\delta_{ML}(X)$ mit
$$\delta_{ML}(x) = \hat{\gamma}_{ML}, \text{ also } L_x(\delta_{ML}(x)) = \max_{\gamma \in \Gamma} L_x(\gamma)$$
heißt *Maximum-Likelihood-Schätzfunktion für* $\gamma \in \Gamma$
(kurz: *ML-Schätzfunktion*)

Beispiel: Normalverteilung

Bei einer einfachen Stichprobe mit Zurücklegen ist:

- $L_x(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi\sigma^2}^n} \cdot e^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}(x_i-\mu)^2}$
- $\ln(\sqrt{2\pi}^n L_x(\mu, \sigma^2)) = -\frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$
- $\frac{\partial}{\partial\mu}(-\frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2) = \frac{1}{2\sigma^2} \sum_{i=1}^{n} 2(x_i - \mu),$

  Nullsetzen ergibt $\widehat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}$

Für die Maximalstelle gilt: $\mu^* = \overline{x} \qquad (*)$

$$\frac{\partial}{\partial(\sigma^2)}\left(-\frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$
$$= -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i-\mu)^2,$$

Nullsetzen und Einsetzen von $\mu^* = \bar{x}$
ergibt $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$
(nicht erwartungstreu)