Beispiel 123

Wir konstruieren mit der ML-Methode einen Schätzer für den Parameter p der Bernoulli-Verteilung. Es gilt $\Pr_p[X_i = 1] = p$ und $\Pr_p[X_i = 0] = 1 - p$. Daraus schließen wir, dass $\Pr_p[X_i = x_i] = p^{x_i}(1-p)^{1-x_i}$, und stellen die Likelihood-Funktion

$$L(\vec{x}; p) = \prod_{i=1}^{n} p^{x_i} \cdot (1-p)^{1-x_i}$$

auf.

Wir suchen als Schätzer für p den Wert, an dem die Funktion L maximal wird. Wir erhalten

$$\begin{aligned}\ln L(\vec{x}; p) &= \sum_{i=1}^{n} (x_i \cdot \ln p + (1 - x_i) \cdot \ln(1-p)) \\ &= n\bar{x} \cdot \ln p + (n - n\bar{x}) \cdot \ln(1-p).\end{aligned}$$

Hierbei bezeichnet $\bar{x}$ das arithmetische Mittel $\frac{1}{n}\sum_{i=1}^{n} x_i$.

Beispiel (Forts.)

Wir finden das Maximum durch Nullsetzen der Ableitung:

$$\frac{d \ln L(\vec{x};p)}{d p} = \frac{n\bar{x}}{p} - \frac{n - n\bar{x}}{1-p} = 0.$$

Diese Gleichung hat die Lösung $p = \bar{x}$.

Beispiel 124

Die Zufallsvariable X sei $\mathcal{N}(\mu, \sigma^2)$-verteilt, und wir suchen Schätzvariablen für die Parameter μ und σ. Nach Definition der Likelihood-Funktion gilt

$$L(\vec{x}; \mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \cdot \prod_{i=1}^{n} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right).$$

Durch Logarithmieren erhalten wir

$$\ln L(\vec{x}; \mu, \sigma^2) = -n(\ln \sqrt{2\pi} + \ln \sigma) + \sum_{i=1}^{n} \left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right).$$

Beispiel 124

Für die Nullstellen der Ableitungen ergibt sich

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^{n} \frac{x_i - \mu}{\sigma^2} \stackrel{!}{=} 0,$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{\sigma^3} \stackrel{!}{=} 0,$$

also

$$\mu = \bar{x} \quad \text{und} \quad \sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2.$$

Wir haben also durch die ML-Methode „fast" das Stichprobenmittel und die Stichprobenvarianz erhalten. Allerdings besitzt der Schätzer für die Varianz hier den Vorfaktor $\frac{1}{n}$ statt $\frac{1}{n-1}$. Die ML-Schätzvariable für die Varianz ist somit nicht erwartungstreu.

3. Konfidenzintervalle

Bei der Verwendung von Schätzvariablen geht man davon aus, dass der erhaltene Schätzwert „nahe“ beim gesuchten Parameter θ liegt. Die Schätzungen werden „besser“, je größer die betrachtete Stichprobe ist. Diese Angaben sind aus quantitativer Sicht natürlich unbefriedigend, da nicht erkennbar ist, wie gut man sich auf den Schätzwert verlassen kann.

Die Lösung dieses Problems besteht darin, statt einer Schätzvariablen U zwei Schätzer U_1 und U_2 zu betrachten. U_1 und U_2 werden so gewählt, dass

$$\Pr[U_1 \leq \theta \leq U_2] \geq 1 - \alpha.$$

Die Wahrscheinlichkeit $1 - \alpha$ heißt Konfidenzniveau und kann dem „Sicherheitsbedürfnis“ angepasst werden.

Wenn wir für eine konkrete Stichprobe die Schätzer U_1 und U_2 berechnen und davon ausgehen, dass $\theta \in [U_1, U_2]$ ist, so ziehen wir höchstens mit Wahrscheinlichkeit α einen falschen Schluss.
$[U_1, U_2]$ heißt Konfidenzintervall.

In vielen Fällen verwendet man nur eine Schätzvariable U und konstruiert mittels $U_1 := U - \delta$ und $U_2 := U + \delta$ ein symmetrisches Konfidenzintervall $[U - \delta, U + \delta]$.

Sei X eine $\mathcal{N}(\mu, \sigma^2)$-verteilte Zufallsvariable, und seien $X_1, \ldots, X_n$ n zugehörige Stichprobenvariablen. Gemäß der Additivität der Normalverteilung (siehe Satz 113) ist das Stichprobenmittel $\overline{X}$ ebenfalls normalverteilt mit $\overline{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$. Wir suchen für $\overline{X}$ ein symmetrisches Konfidenzintervall.

Nach Satz 99 ist

$$Z := \sqrt{n} \cdot \frac{\overline{X} - \mu}{\sigma}$$

standardnormalverteilt.

Für Z betrachten wir das Konfidenzintervall $[-c, c]$ für ein geeignetes $c > 0$ und setzen

$$\Pr[-c \leq Z \leq c] \stackrel{!}{=} 1 - \alpha.$$

Auflösen nach μ ergibt

$$\Pr\left[\overline{X} - \frac{c\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + \frac{c\sigma}{\sqrt{n}}\right] \stackrel{!}{=} 1 - \alpha\,.$$

Das gesuchte Konfidenzintervall lautet also

$$K = [\overline{X} - \frac{c\sigma}{\sqrt{n}}, \overline{X} + \frac{c\sigma}{\sqrt{n}}]\,.$$

Den Parameter c wählen wir wie folgt:

$$\Pr[-c \leq Z \leq c] = \Phi(c) - \Phi(-c) \stackrel{!}{=} 1 - \alpha.$$

Wegen der Symmetrie von Φ gilt $\Phi(-x) = 1 - \Phi(x)$ und wir erhalten

$$\Phi(c) - \Phi(-c) = 2 \cdot \Phi(c) - 1 \stackrel{!}{=} 1 - \alpha \iff \Phi(c) = 1 - \frac{\alpha}{2},$$

also

$$c = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right).$$

Definition 125

X sei eine stetige Zufallsvariable mit Verteilung F_X. Eine Zahl x_γ mit

$$F_X(x_\gamma) = \gamma$$

heißt γ-Quantil von X bzw. der Verteilung F_X.

Definition 126

Für die Standardnormalverteilung bezeichnet z_γ das γ-Quantil.

Damit können wir das gesuchte Konfidenzintervall angeben durch

$$K = \left[\overline{X} - \frac{z_{(1-\frac{\alpha}{2})}\sigma}{\sqrt{n}}, \overline{X} + \frac{z_{(1-\frac{\alpha}{2})}\sigma}{\sqrt{n}}\right].$$