WS 2009/10

Diskrete Strukturen

Dr. Werner Meixner

Fakultät für Informatik
TU München

http://www14.in.tum.de/lehre/2009WS/ds/uebung/

1. Juli 2010

ZÜ X Vorbereitung Blatt 10

1. VA 1

Eine Werbeagentur möchte am letzten Tag der Fußballweltmeisterschaft mit einer Blitzumfrage schätzen, welcher Anteil ϑ der per Bahn anreisenden Fußballfans einen Platz im Stadion hat.

Jeder der 12 Mitarbeiter befragt so lange zufällig ausgewählte Fans, bis er einen Fan gefunden hat,
der eine Karte für das Stadion besitzt.

Die Anzahl der vom Mitarbeiter i befragten Fans sei X_i.

1. Man bestimme auf der Basis der ermittelten Stichprobenwerte

$$3,\ 1,\ 2,\ 3,\ 2,\ 1,\ 2,\ 1,\ 5,\ 3,\ 2,\ 3$$

einen Maximum-Likelihood-Schätzwert für ϑ.

Lösung:

Für die Likelihood-Funktion $L(\vec{x};\vartheta)$
der Stichprobe $\vec{x} = (x_1, \ldots, x_{12})$ gilt mit $n = 12$

$$\begin{aligned} L(\vec{x};\vartheta) &= \prod_{i=1}^{n} \Pr_\vartheta[X_i = x_i] \\ &= \prod_{i=1}^{n} \left[(1-\vartheta)^{x_i - 1} \cdot \vartheta\right]. \end{aligned}$$

Wir schreiben $\sum_{i=1}^{n} x_i - n = n\bar{x} - n$ und erhalten

$$L(\vec{x};\vartheta) = (1-\vartheta)^{n\bar{x}-n} \cdot \vartheta^n .$$

Gesucht ist der Wert $\vartheta = \bar{\vartheta}$ mit $0 \le \vartheta \le 1$, für den $L(\vec{x};\vartheta)$ das Maximum $L(\vec{x};\bar{\vartheta})$ in dem $[0,1]$-Intervall annimmt.

Falls $\bar{x} = 1$,
dann folgt sofort $\vartheta = 1$.

Dies bedeutet, dass die Annahme $\vartheta = 1$ die maximale Wahrscheinlichkeit für das Testergebnis liefert,
d.h. hier, dass jeder Fan eine Stadionkarte besitzt.

Falls $\bar{x} \neq 1$,
dann bestimmen wir das Maximum von ϑ zwischen 0 und 1 mithilfe der Nullstelle der Ableitung von L nach ϑ wie folgt.

$$
\begin{aligned}
\frac{\mathsf{d}\, L(\vec{x};\vartheta)}{\mathsf{d}\,\vartheta} &= (n\bar{x}-n)(1-\vartheta)^{n\bar{x}-n-1}(-1)\vartheta^{n} + (1-\vartheta)^{n\bar{x}-n} n\vartheta^{n-1}\\
&= (1-\vartheta)^{n\bar{x}-n-1}\vartheta^{n-1}[(n\bar{x}-n)(-\vartheta)+n(1-\vartheta)]\\
&= (1-\vartheta)^{n\bar{x}-n-1}\vartheta^{n-1}[n-n\bar{x}\vartheta]
\end{aligned}
$$

Als Nullstellen der Ableitung der Likelihood-Funktion erhalten wir

$$
\vartheta = 0, \quad \vartheta = 1 \quad \text{und} \quad \vartheta = \frac{1}{\bar{x}}\,.
$$

Man beachte, dass $0 < \frac{1}{\bar{x}} < 1$ gilt.

Wegen

$$L(\vec{x};0) = L(\vec{x};1) = 0 \quad \text{und} \quad L\left(\vec{x};\frac{1}{\bar{x}}\right) > 0$$

scheiden $\vartheta = 0$ und $\vartheta = 1$ als Maximumstellen aus.

Es folgt

$$\bar{\vartheta} = \frac{1}{\bar{x}}\,.$$

Nun berechnen wir $\frac{1}{\bar{x}}$ aus der Stichprobe mit $\bar{x} = \frac{7}{3}$ und erhalten

$$\bar{\vartheta} = \frac{3}{7}\,.$$

Ergebnis:

Die Annahme, dass $\frac{3}{7}$ der Fans Stadionkarten besitzen,
verleiht dem Testergebnis die maximale Wahrscheinlichkeit.

1. Man gebe mit Hilfe des Zentralen Grenzwertsatzes ein konkretes 95%-Konfidenzintervall für ϑ an.

   Das heißt:

   Es sind Schranken α_1, α_2 zu konstruieren, so dass gilt

$$\Pr[\alpha_1 \leq \bar{\vartheta} \leq \alpha_2] \geq 0,95\,.$$

Lösung:

Die Stichprobe liefert den Wert $\bar{x} = \frac{7}{3}$ für die Zufallsvariable $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ mit $n = 12$.

Für das Konfidenzintervall für ϑ
zum Konfidenzniveau $1 - \alpha = 0.95$
setzen wir an

$$\Pr\left[-c \leq \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} \leq c\right] \approx 0.95$$

mit dem Quantil $c = x_{1-\frac{\alpha}{2}} = x_{0.975} \approx 1.96$

und den Gleichungen $\mu_{\bar{X}} = \frac{1}{\vartheta}$ und $\sigma_{\bar{X}}^2 = \frac{1-\vartheta}{n\vartheta^2}$ für $n = 12$.

Wir bestimmen nun die
Menge aller ϑ, für die die Ungleichung $-c \leq \frac{\bar{X}-\mu_{\bar{X}}}{\sigma_{\bar{X}}} \leq c$ gilt.

$$
\begin{aligned}
-c \leq \frac{\bar{X}-\mu_{\bar{X}}}{\sigma_{\bar{X}}} \leq c \quad & \iff \quad (\bar{X}-\mu_{\bar{X}})^2 \leq c^2\sigma_{\bar{X}}^2 \\
& \iff \quad (\bar{X}-\frac{1}{\vartheta})^2 \leq c^2 \cdot \frac{1}{n} \cdot \frac{1-\vartheta}{\vartheta^2} \\
& \iff \quad (\vartheta\bar{X}-1)^2 \leq \frac{c^2}{n} \cdot (1-\vartheta) \\
& \iff \quad \vartheta^2\bar{X}^2 + (\frac{c^2}{n} - 2\bar{X})\vartheta + (1-\frac{c^2}{n}) \leq 0
\end{aligned}
$$

Nun setzen wir sämtliche Zahlenwerte ein und erhalten
(wegen $c \approx 1.96$ näherungsweise)

$$-c \leq \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} \leq c \iff 49\vartheta^2 - 39.1188\vartheta + 6.1188 \leq 0\,.$$

Lösung wie folgt.

Nullstellen der quadratischen Gleichung nach bekannter Formel:

$$49\vartheta^2 - 39.1188\vartheta + 6.1188 = 0\,.$$

Nullstellen seien α_1, α_2, wobei wir $\alpha_1 < \alpha_2$ annehmen können. Dann ergibt sich als Lösung

$$\alpha_1 \leq \vartheta \leq \alpha_2$$

mit $\alpha_1 \approx 0.2135$ und $\alpha_2 \approx 0.5848\,.$

Ergebnis:

$$\Pr[0.2135 \leq \bar{\vartheta} \leq 0.5848] \geq 0,95\,.$$