# Diskrete Strukturen II

## Zentralübung

Hanjo Täubig Sebastian Wernicke

09.07.2004

# Raumänderung

Die Zentralübung findet an folgenden Tagen nicht im gewohnten Hörsaal statt, sondern wie folgt:

- Freitag, 16.07.2004: MW 1801
- Freitag, 23.07.2004: ZÜ entfällt,
  stattdessen:
  Sprechstunde am Mittwoch, 21.07.2004
  ab 9:00 Uhr in Raum 03.09.011

# Aufgabe 1

Beim Testen von Hypothesen bezeichnet man die zu überprüfende Hypothese (*Nullhypothese*) generell mit $H_0$ und die *Alternative* mit $H_1$.

Ein Tierhändler erhält ein Paket mit 100 Frettchen.
Er will testen, ob weniger als zehn ($< 10$) dieser Frettchen aggressiv und bissig sind. Dazu hält er zehn Frettchen seinen Finger hin und nimmt das Paket nur an, wenn ihn keins davon beißt (wir nehmen an, dass ein aggressives Frettchen sofort zubeißen würde).

Wie lauten die Hypothesen des Händlers?
Was ist das Niveau des Tests?

# Aufgabe 1

$X$: Anzahl der aggressiven Frettchen in der 10er-Stichprobe
$S$: Anzahl der aggressiven Frettchen im gesamten Paket

$X$ ist hypergeometrisch verteilt mit $N = 100$ und $S$.

Der Händler testet die Hypothese

$$H_0 : S \leq 9$$

gegen

$$H_1 : S \geq 10$$

und verwirft die Nullhypothese dabei im Fall $X \geq 1$.

# Aufgabe 1

Für $S = 9$ ist

$$\Pr_S[X > 0] = 1 - \Pr_S[X = 0] = 1 - \frac{\binom{9}{0}\binom{91}{10}}{\binom{100}{10}} = 62.9\%.$$

Das ist auch das effektive Signifikanzniveau des Tests, denn für $S < 9$ ist die Wahrscheinlichkeit, dass die Nullhypothese fälschlicherweise verworfen wird, kleiner.

# Aufgabe 2

In einer großen Stadt gibt es $N$ Taxis, die mit den Nummern $1, \ldots, N$ beschriftet sind.
Wir stehen an der Straße und beobachten die Taxis, dabei notieren wir uns deren Nummer. (Wiederholungen werden ignoriert).
Sei $x_1, \ldots, x_n$ die aufsteigend sortierte Folge der beobachteten Nummern. Nun wollen wir die Anzahl $N$ der Taxis schätzen.

a) Was ist der Maximum Likelihood Schätzer für $N$?
Ist dieser erwartungstreu?

b) Geben Sie einen Schätzer für die Anzahl der Taxis an, der $N$ dadurch abschätzt, dass er versucht, die Größe der nicht beobachteten Lücke $x_n + 1, \ldots N$ oberhalb von $x_n$ abzuschätzen.

# Aufgabe 2

a) Es gilt offensichtlicherweise $N \geq x_n$.

Die Wahrscheinlichkeit $\Pr_N[\{x_1, \ldots x_n\}]$ ("Wahrscheinlichkeit, $\{x_1, \ldots x_n\}$ zu beobachten, wenn es $N$ Taxis gibt") ist $\binom{N}{n}^{-1}$, denn bei $N$ Taxis ist jede Teilmenge der Mächtigkeit $n$ gleichwahrscheinlich.

Für einen Maximum-Likelihood Schätzer gilt es, $\binom{N}{n}^{-1}$ in Abhängigkeit von $N$ zu maximieren, was genau für $\widehat{N}(\{x_1, \ldots x_n\}) = x_n$ der Fall ist, man schätzt also die Anzahl der Taxis durch die höchste beobachtete Nummer ab.

Dieser Schätzer ist offensichtlich nicht erwartungstreu, weil man oft zu niedrig, aber nie zu hoch schätzt.

# Aufgabe 2

b) Wir wählen nun folgenden Ansatz:
Zwischen den beobachteten $x_i$ haben wir "Lücken".

Wir versuchen nun, dadurch die Größe der Lücke von $x_n$ zu $N$ durch die mittlere Lückengröße in $\{x_1, \dots x_n\}$ abzuschätzen.

Die mittlere Lückengröße ist

$$\frac{1}{n}((x_1 - 1) + (x_2 - x_1) + \cdots + (x_n - x_{n-1})) = \frac{x_n - 1}{n}$$

und führt zum Schätzer

$$N_1(\{x_1, \dots x_n\}) = x_n + \frac{x_n - 1}{n}.$$

# Aufgabe 3

Sei $X$ die Anzahl der Unfälle in einer Stadt pro Woche und Poisson-verteilt mit Parameter $\lambda > 0$.
Wir wollen aus der Beobachtung von $X$ die Wahrscheinlichkeit ableiten, dass in den folgenden drei Wochen kein Unfall geschieht (also $(\Pr[X = 0])^3$).

Zeigen Sie: Ist dieser Schätzer erwartungstreu, dann liefert er unsinnige Schätzwerte!

Hinweis: $\sum_{i=0}^{\infty} k^i \cdot \frac{\lambda^i}{i!} = e^{k\cdot\lambda}$ kann dazu verwendet werden, $k$ zu bestimmen.

# Aufgabe 3

Für einen erwartungstreuen Schätzer $T(x)$ müsste unabhängig von $\lambda$ gelten:

$$\begin{aligned} \mathbb{E}(T(x)) &= \left(e^{-\lambda} \cdot \frac{\lambda^0}{0!}\right)^3 = e^{-3\lambda} \\ \Rightarrow \quad \sum_{i=0}^{\infty} T(i) \cdot e^{-\lambda} \frac{\lambda^i}{i!} &= e^{-3\lambda} \end{aligned}$$

Mit Hilfe des Hinweises wissen wir, dass daraus unmittelbar

$$T(x) = (-2)^x$$

folgt, was unsinnig ist, da dieser Schätzer sowohl Werte $< 0$ als auch $> 1$ produziert.

# Aufgabe 4

Seien $X_1, X_2, \dots X_n$ unabhängig und gleichverteilt in $\{1, 2, \dots b\}$, $b \in \mathbb{N}$.

Wenn $X = \max\{X_1, \dots, X_n\}$, geben Sie ein Konfidenzintervall für $b$ zum Niveau $1 - \alpha$ aufgrund der Beobachtung von $X$ an.

Hinweis: Hier handelt es sich beim Konfidenzintervall um eine endliche Menge.

# Aufgabe 4

Wegen $\Pr[b < x] = 0$ suchen wir eine Menge

$$C(x) = \{x, x+1, \ldots, b^*(x)\},$$

welche möglichst klein ist, jedoch so viele Werte enthält, dass

$$\Pr_b[b \in C(x)] \geq 1 - \alpha \quad \Longrightarrow \quad \Pr_b[b \notin C(x)] \leq \alpha$$

gilt.

Also ist $b^*(x)$ die kleinstmögliche Zahl $> x$, für welche

$$\Pr_b[b^*(X) \leq b] \leq \alpha$$

gilt.

# Aufgabe 4

Wegen

$$Pr_b[X \le x] = \left(\frac{x}{b}\right)^n$$

folgt unmittelbar

$$\begin{aligned} b^*(x) &= \min\left\{ i \mid \left(\frac{x}{i}\right)^n \le \alpha \right\} \\ &= \min\left\{ i \mid \left(\frac{x}{i}\right) \le \sqrt[n]{\alpha} \right\} \\ &= \min\left\{ i \mid \left(\frac{1}{i}\right) \le \frac{\sqrt[n]{\alpha}}{x} \right\} \\ &= \min\left\{ i \mid \quad i \quad \ge \frac{x}{\sqrt[n]{\alpha}} \right\} \end{aligned}$$

Also ist

$$C(x) = \left\{ x, x+1, \ldots, \left\lceil \frac{x}{\sqrt[n]{\alpha}} \right\rceil \right\}.$$

# Raumänderung

Die Zentralübung findet an folgenden Tagen nicht im gewohnten Hörsaal statt, sondern wie folgt:

- Freitag, 16.07.2004: MW 1801
- Freitag, 23.07.2004: ZÜ entfällt,
  stattdessen:
  Sprechstunde am Mittwoch, 21.07.2004
  ab 9:00 Uhr in Raum 03.09.011