# 5 Aufgaben zur Punktschätzung

**Zoltán Zomotor**

**Versionsstand: 20. Mai 2015, 10:04**



### Aufgabe 1: Reviewfragen

(a) Wozu dient die Punktschätzung?

(b) Was bedeutet *erwartungstreu* und *asymptotisch erwartungstreu*?

(c) Wozu dient die Bestimmung der *Wirksamkeit* einer Schätzufunktion?

(d) Was bedeutet hier der Begriff *Konsistenz*?

(e) Was ist der Unterschied zwischen einem Maximum-Likelihood-Schätzwert und einer Maximum-Likelihood-Schätzfunktion

(f) $\hat{\Theta} = f(X_1, X_2, \ldots, X_n)$ sei eine Maximum-Likelihood-Schätzfunktion für den Parameter $\vartheta$. Unter welcher Voraussetzung ist $\sin \hat{\Theta}$ eine Maximum-Likelihood-Schätzfunktion von $\sin \vartheta$?

### Aufgabe 2:

Der Erwartungswert $\mu$ in der Grundgesamtheit soll durch folgende Stichprobenfunktionen geschätzt werden:

$$\hat{\Theta}_n = \frac{1}{n} \cdot (X_1 + X_2 + \cdots + X_n), \quad \hat{\Theta}'_n = \frac{2}{n^2} \sum_{i=1}^{n} i X_i, \quad \hat{\Theta}''_n = \frac{2}{n(n+1)} \cdot (X_1 + 2X_2 + \cdot + nX_n)$$

Prüfen Sie die Erwartungstreue der Schätzfunktionen und ermitteln Sie die wirksamste unter den erwartungstreuen Schätzfunktionen. Sind die Schätzfunktionen konsistent?
Hinweis:

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2} \quad \text{und} \quad \sum_{i=1}^{n} i^2 = \frac{1}{3} n(n+1)(n+\frac{1}{2})$$

**Aufgabe 3:**

Es sei $x_1, \ldots, x_n$ das Ergebnis einer einfachen Stichprobe aus einer Poisson-verteilten Grundgesamtheit mit dem unbekannten Parameter $\mu$.

a) Bestimmen Sie die Likelihood-Funktion für dieses Stichprobenergebnis

b) Zeigen Sie, dass $\bar{X}$ die Maximum-Likelihood-Schätzfunktion für den unbekannten Parameter $\mu$ ist.

c) Bestimmen Sie die Maximum-Likelihood-Schätzfunktion für die Standardabweichung $\sqrt{\mu}$.

**Aufgabe 4:**

In einem gut gefüllten Saal befinden sich $N$ Personen, die mit 1 bis $N$ durchnummeriert sind. Herr Maier von der Gewerbeaufsicht kennt $N$ nicht und wird vom Veranstalter mit allerlei Tricks am Betreten des Saals gehindert. Herr Maier, der $N$ ermitteln möchte, ist lediglich in der Lage, die Nummer eines Teilnehmers, der zufällig den Saal verlässt, zu erkennen. Fassen Sie diese Nummer als Realisation $x$ einer Zufallsvariablen $X$ auf.

a) Bestimmen Sie die Wahrscheinlichkeitsfunktion von $X$ (in Abhängigkeit von $N$).

b) Ist $2X - 1$ eine erwartungstreue Schätzfunktion für $N$?

**Aufgabe 5:**

$X_1, \ldots, X_n$ und $Y_1, \ldots, Y_n$ seien einfache Stichproben, wobei die Verteilung der $X_i, Y_i$ vom unbekannten Parameter $p \in [0, \frac{1}{2}]$ abhängt. Für alle $i = 1, \ldots, n$ gelte:

$$\begin{aligned}
\mathrm{E}(X_i) &= \mathrm{E}(X_i^2) = p \\
\mathrm{E}(Y_i) &= 1.5p \\
\mathrm{Var}(Y_i) &= 1.5p(1 - 1.5p) \\
\mathrm{E}(X_i \cdot Y_i) &= p^2
\end{aligned}$$

a) Berechnen Sie in Abhängigkeit von $p$ die Größen

$$\mathrm{Var}(X_i), \quad \mathrm{Cov}(X_i, Y_i), \quad \mathrm{E}(X_i + Y_i) \text{ und } \mathrm{Var}(X_i + Y_i)$$

b) Bestimmen Sie unter den drei Schätzfunktionen

$$\begin{aligned}
\hat{\Theta}_1 &= \frac{1}{n} \sum_{i=1}^{n} X_i \\
\hat{\Theta}_2 &= \frac{1}{n} \sum_{i=1}^{n} Y_i \\
\hat{\Theta}_3 &= \frac{2}{5n} \sum_{i=1}^{n} (X_i + Y_i)
\end{aligned}$$

die für $p$ erwartungstreuen, und bestimmen Sie unter den erwartungstreuen die wirksamste.

**Aufgabe 6:**

Zwei unterschiedliche Messgeräte $A$ und $B$ messen laufend den Salzgehalt in einem Meerwasseraquarium. Beide Messgeräte liefern „im Mittel richtige“ Werte, was sich durch

$$E(X) = E(Y) = E(S) = \mu$$

beschreiben lässt, wobei $X$ (bzw. $Y$) das vom Messgerät $A$ (bzw. $B$) verwendete Messverfahren zur Messung des Salzgehalts $S$ ist. Es ist bekannt, dass $A$ ungenauere Messungen als $B$ liefert, was sich durch

$$\mathrm{Var}(X) = 1.5 \cdot \mathrm{Var}(Y)$$

beschreiben lässt. Ferner sind $X$ und $Y$ unkorreliert. Beide Messgeräte liefern ihre Messergebnisse an einen Zentralrechner, der aus $X$ und $Y$ mittels den Gewichten $a$ und $b$ die Schätzfunktion

$$Z = a \cdot X + b \cdot Y$$

bildet.

(a) Wie müssen $a$ und $b$ gewählt werden, damit $Z$ erwartungstreu für $\mu$ ist?

(b) Welche unter den erwartungstreuen Schätzfunktionen $Z$ ist die wirksamste?

**Aufgabe 7:**

Gegeben sei eine einfache Stichprobe $X_1, \ldots, X_n$ aus einer Grundgesamtheit mit unbekanntem Erwartungswert $\mu$, der mit folgender Stichprobenfunktion geschätzt werden soll

$$\hat{\Theta} = \frac{1}{n-1} \sum_{i=1}^{n-1} X_i + \alpha \cdot X_n \quad \text{mit } \alpha \in [-1; 1].$$

(a) Bestimmen Sie $\alpha$ so, dass $\hat{\Theta}$ erwartungstreu für $\mu$ ist. Ist die resultierende Schätzfunktion konsistent?

(b) Begründen Sie, ob das die wirksamste Schätzfunktion für $\mu$ ist.

(c) Angenommen, die Grundgesamtheit ist $N(\mu, 9)$ verteilt. Wie hoch ist dann die Wahrscheinlichkeit, dass der von $\hat{\Theta}$ gelieferte Schätzwert $\hat{\vartheta}$ bei einem Stichprobenumfang von $n = 82$ im Intervall $\mu - 1; \mu + 1$ liegen wird?

**Aufgabe 8:**

Eine einfache Stichprobe bezüglich eines Merkmals $X$ in einer Grundgesamtheit ergab folgende Ergebnisse:

| $x$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $h(x)$ | 3 | 4 | 4 | 3 | 2 | 1 |

Nehmen Sie zunächst an, $X$ sei exponentialverteilt.

(a) Bestimmen Sie den Maximum-Likelihood-Schätzwert für $\lambda$.

(b) Wie groß ist unter Verwendung des Ergebnisses aus (a) die Wahrscheinlichkeit, dass $X$ einen Wert annimmt, der nicht kleiner als 2 ist?

(c) Lösen Sie die Teile (a) und (b) nun unter der Annahme, dass $X$ Poisson-verteilt ist.

(d) Bei dem Merkmal $X$ handelt es sich um die Anzahl der zwischen 23:00 und 24:00 Uhr an einer Tankstelle tankenden Kraftfahrzeuge. Begründen Sie kurz, welche der beiden obigen Verteilungsannahmen sinnvoller ist.

**Aufgabe 9:**

Die Zeitdauer $T$ für einen bestimmten Vorgang sei gemäß der Verteilungsfunktion

$$F(t) = 1 - \mathrm{e}^{-b\sqrt{t}} \text{ für } t > 0$$

verteilt (mit $F(t) = 0$ für $t \leq 0$).

(a) Bestimmen Sie den Median $t_{\text{Med}}$ in Abhängigkeit vom Parameter $b$.

(b) Bei einer einfachen Stichprobe zum Merkmal $T$ vom Umfang $n = 10$ wurden die Zeitdauern

$$(1, 4, 9, 9, 4, 4, 4, 1, 1, 4)$$

registriert. Berechnen Sie den Maximum-Likelihood-Schätzwert $\hat{b}$ für $b$. (Die zweite Ableitung muss nicht ausgewertet werden.)

(c) Berechnen Sie den Maximum-Likelihood-Schätzwert $\hat{t}_{\text{Med}}$ für den Median aus der in (b) angegebenen Stichprobenrealisation.

**Aufgabe 10:**

Freikletterer Hans interessiert, wie groß die Entfernung $Z$ von seiner frei werdenden Hand zu einem zufällig gewählten Zeitpunkt während des Kletterns zum nächsten „Henkel“ (großer guter Griff für alle Finger) ist. Er geht davon aus, dass $Z$ gleichverteilt über dem Intervall $[0, d]$ ist.

(a) Stellen Sie die Likelihood-Funktion $f(z_1, z_2, \ldots, z_n|d)$ für eine Stichprobenrealisation $(z_1, z_2, \ldots, z_n)$ auf.

(b) Für $n = 5$ Entfernungsmessungen (in Armlängen) hat er sich die Stichprobenrealisation

$$(z_1, \ldots, z_5) = (6, 18, 11, 20, 9)$$

gemerkt. Bestimmen Sie den Maximum-Likelihood-Schätzwert $\hat{d}$ für $d$.

**Lösung 1:**

(a) Schätzung eines unbekannten Parameters $\theta$ auf Basis einer Stichprobe.

(b) $\mathrm{E}(\hat{\Theta}) = \vartheta$, $\lim\limits_{n\to\infty} \mathrm{E}(\hat{\Theta}) = \theta$

(c) Vergleich zweier Schätzfunktionen aufgrund ihrer Varianz. Die mit der kleineren Varianz ist wirksamer und damit besser.

(d) $\mathrm{E}(\hat{\Theta}_n) = \vartheta$ und $\lim\limits_{n\to\infty} \mathrm{Var}()\hat{\Theta}_n = 0$

(e) ML-Schätzwert: Maximierung der Likelihoodfunktion für ein konkretes Stichprobenergebnis.
ML-Schätzfunktion: Maximierung der Likelihoodfunktion für allgmeine Stichprobe.

(f) $X_i > 0$

**Lösung 2:**

$$\begin{aligned}
E(\hat{\Theta}_n) &= E(\frac{1}{n} \cdot (X_1 + X_2 + \cdots + X_n)) \\
&= \frac{1}{n} \cdot (E(X_1) + E(X_2) + \cdots + E(X_n)) = \frac{1}{n} \cdot (\mu + \mu + \cdots + \mu) = \frac{1}{n} \cdot (n\mu) \\
&= \mu \Rightarrow \quad \text{erwartungstreu.} \\
\mathrm{Var}(\hat{\Theta}_n) &= \mathrm{Var}(\frac{1}{n} \cdot (X_1 + X_2 + \cdots + X_n)) \\
&= \frac{1}{n^2} \cdot (\mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_n)) = \frac{1}{n^2} \cdot (n\sigma^2) \\
&= \frac{\sigma^2}{n}
\end{aligned}$$

$$\begin{aligned}
E(\hat{\Theta}'_n) &= E(\frac{2}{n^2} \sum_{i=1}^{n} iX_i) = \frac{2}{n^2} \sum_{i=1}^{n} iE(X_i) = \frac{2}{n^2} \mu \sum_{i=1}^{n} i \\
&= \frac{2\mu}{n^2} \cdot \frac{n(n+1)}{2} = \mu(1 + \frac{1}{n}) \Rightarrow \ \textit{nicht} \text{ erwartungstreu und } \textit{nicht} \text{ konsistent}
\end{aligned}$$

$$
\begin{aligned}
E(\hat{\Theta}_n'') &= E\Big(\frac{2}{n(n+1)} \cdot (X_1 + 2X_2 + \cdot + nX_n)\Big) = \frac{2}{n(n+1)} \cdot \Big(E(X_1) + 2E(X_2) + \cdot + nE(X_n)\Big) \\
&= \frac{2}{n(n+1)}\, \mu\, \frac{n(n+1)}{2} \\
&= \mu \Rightarrow \quad \text{erwartungstreu.}
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{Var}(\hat{\Theta}_n'') &= \operatorname{Var}\Big(\frac{2}{n(n+1)} \cdot (X_1 + 2X_2 + \cdot + nX_n)\Big) = \left(\frac{2}{n(n+1)}\right)^2 \cdot (1^2 \operatorname{Var}(X_1) + 2^2 \operatorname{Var}(X_2) + \cdot + n^2 \mathrm{V} \\
&= \left(\frac{2}{n(n+1)}\right)^2 \cdot \frac{1}{3} n(n+1)(n+\frac{1}{2}) \cdot \sigma^2 \\
&= \frac{4(n+\frac{1}{2})}{3n(n+1)} \sigma^2
\end{aligned}
$$

Wirksamkeit:

$$
\begin{aligned}
\operatorname{Var}(\hat{\Theta}_n'') &> \operatorname{Var}(\hat{\Theta}_n) \\
\Leftrightarrow \frac{4(n+\frac{1}{2})}{3n(n+1)} \sigma^2 &> \frac{\sigma^2}{n} \\
\Leftrightarrow \frac{4}{3} \frac{1 + \frac{1}{2n}}{1 + \frac{1}{n}} &> 1
\end{aligned}
$$

ist für $n \to \infty$ gegeben, d.h. $\hat{\Theta}_n$ ist die wirksamste Schätzfunktion.

**Lösung 3:**

a) $f(x_1, \ldots, x_n|\mu) = \dfrac{\mu^{\sum_{i=1}^n x_i} \mathrm{e}^{-n\mu}}{\prod_{i=1}^n x_i!} = \dfrac{\mu^{n\bar{x}} \mathrm{e}^{-n\mu}}{\prod_{i=1}^n x_i!}$

b)

$$
\begin{aligned}
\ln f(x_1, \ldots, x_n|\mu) &= n\bar{x} \ln \mu - n\mu - \prod_{i=1}^n x_i! \\
\frac{\mathrm{d} \ln f}{\mathrm{d}\mu} &= \frac{n\bar{x}}{\mu} - n \stackrel{!}{=} 0 \\
&\Rightarrow \hat{\mu} = \bar{x}
\end{aligned}
$$

c) $s = \sqrt{\hat{\mu}} = \sqrt{\bar{x}}$

**Lösung 4:**

(a) $f(x) = \begin{cases} \dfrac{1}{N}, & \text{für } x = 1, \ldots, N \\ 0, & \text{sonst} \end{cases}$

(b) Ja, weil:

$$\mathrm{E}(2X-1) = 2\,\mathrm{E}(X) - 1 = 2\sum_{i=1}^{N} i f(i) - 1 = \frac{2}{N}\cdot\frac{N\,(N+1)}{2} - 1 = N$$

**Lösung 5:**

(a)

$$\mathrm{Var}(X_i) = p\,(1-p) \quad \mathrm{Cov}(X_i, Y_i) = 0.5p^2 \quad \mathrm{E}(X_i + Y_i) = 2.5p$$
$$\mathrm{Var}(X_i + Yi) = p\,(1-p) + 1.5p(1-1.5p) - 2\cdot 0.5p^2 = 2.5p - 4.25p^2$$

(b)

$$\mathrm{E}(\hat{\Theta}_1) = \mathrm{E}(X_i) = p \Rightarrow \hat{\Theta}_1 \text{ ist erwartungstreu für } p$$
$$\mathrm{E}(\hat{\Theta}_2) = \mathrm{E}(Y_i) = 1.5p \Rightarrow \hat{\Theta}_2 \text{ ist nicht erwartungstreu für } p$$
$$\mathrm{E}(\hat{\Theta}_3) = \frac{2}{5}\Big(\mathrm{E}(X_i) + \mathrm{E}(Yi)\Big) = p \Rightarrow \hat{\Theta}_3 \text{ ist erwartungstreu für } p$$
$$\mathrm{Var}(\hat{\Theta}_1) = \frac{1}{n}p\,(1-p)$$
$$\mathrm{Var}(\hat{\Theta}_3) = \frac{0.4}{n}p\,(1-1.7p)$$
$$\mathrm{Var}(\hat{\Theta}_1) \geq \mathrm{Var}(\hat{\Theta}_3) \Leftrightarrow \frac{1}{n}p\,(1-p) \geq \frac{0.4}{n}p\,(1-1.7p) \Rightarrow p \leq 1.875$$

Wegen $p \in [0, \frac{1}{2}]$ ist $p \leq 1.875$ immer erfüllt $\Rightarrow$ $\hat{\Theta}_3$ ist wirksamer als $\hat{\Theta}_3$

**Lösung 6:**

(a) $\mathrm{E}(Z) = a\mu + b\mu = (a+b)\mu \overset{!}{=} \mu \quad \Leftrightarrow \quad a+b=1$

(b) Minimiere die Varianz von $Z = aX + (1-a)Y$ bezüglich $a$:

$$\frac{\mathrm{d}\,\mathrm{Var}(Z)}{\mathrm{d}a} = 3a\cdot\mathrm{Var}(Y) - 2(1-a)\cdot\mathrm{Var}(Y) \overset{!}{=} 0 \Rightarrow a = \frac{2}{5}, b = \frac{3}{5}$$
$$\frac{\mathrm{d}^2\,\mathrm{Var}(Z)}{\mathrm{d}a^2} = 5\cdot\mathrm{Var}(Y) > 0, \text{ also wurde ein Minimum erreicht}$$

**Lösung 7:**

(a) $\mathrm{E}\,\hat{\Theta}) = \mu\cdot(1+\alpha) \overset{!}{=} \mu \quad \Leftrightarrow \quad \alpha = 0$
dann: $\mathrm{Var}(\hat{\Theta}) = \dfrac{\mathrm{Var}(X_i)}{n-1} \to 0$ für $n \to \infty \Rightarrow \hat{\Theta}$ ist konsistent

(b) Nein, denn $\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ ist wirksamer.

(c) $\mathrm{Var}(\hat{\Theta}) = \frac{9^2}{82-1} = 1 \Rightarrow \hat{\Theta} \sim N(\mu, 1) \Rightarrow P(\mu - 1 < \hat{\Theta} < \mu + 1) = 2\Phi(1) - 1 = 0.683$

**Lösung 8:**

(a) $\hat{\lambda} = \frac{1}{\bar{x}}$, $\bar{x} = 2 \Rightarrow \hat{\lambda} = 0.5$

(b) stetige Zufallsvariable: $P(X \geq 2) = 1 - P(X \leq 2) = 1 - (1 - \mathrm{e}^{-0.5 \cdot 2}) = 0.368$

(c) $\hat{\lambda} = \bar{x} = 2$
diskrete Zufallsvariable: $P(X \geq 2) = 1 - F(1) = 1 - 0.406 = 0.594$

(d) Die Poisson-Verteilung ist sinnvoller, weil $X$ hier eine diskrete Zufallsvariable ist. Außerdem werden hier seltene Ereignisse betrachtet. Die Exponentialverteilung ist dagegen eher zur Modellierung von Wartezeiten oder Lebensdauern geeignet.

**Lösung 9:**

(a) $F(_{\mathrm{Med}}) = 1 - \mathrm{e}^{-b\sqrt{t_{\mathrm{Med}}}} = 0.5 \Rightarrow t_{\mathrm{Med}} = \left(\frac{\ln 2}{b}\right)^2$

(b)

$$\begin{aligned}
f(t) &= \frac{\mathrm{d}F(t)}{\mathrm{d}t} = \frac{b\mathrm{e}^{-b\sqrt{t}}}{2\sqrt{t}} \\
f(1, \ldots, 4|b) &= \frac{b^{10} \cdot \mathrm{e}^{-19b}}{2^{10} \cdot \sqrt{1} \cdot \cdots \cdot \sqrt{4}} \\
\ln f(1, \ldots, 4|b) &= 10 \ln b - 19b - \ln(2^{10} \cdot \sqrt{1} \cdot \cdots \cdot \sqrt{4}) \\
\frac{\partial}{\partial b} \ln f(1, \ldots, 4|b) &= \frac{10}{b} - 19 \stackrel{!}{=} 0 \\
\hat{b} &= \frac{10}{19} = 0.526
\end{aligned}$$

(c) $\hat{t}_{\mathrm{Med}} = \left(\frac{\ln 2}{0.526}\right)^2 = 1.737$

**Lösung 10:**

(a) Dichtefunktion der Gleichverteilung über $[0, a]$:

$$f(z) = \begin{cases} \frac{1}{a}, & \text{für } 0 \leq z \leq a \\ 0, & \text{sonst} \end{cases}$$

$\Rightarrow$ LikelihoodFunktion:

$$f(z_1, \ldots, z_n|a) = \begin{cases} \frac{1}{a^n}, & \text{falls } \min\limits_i z_i \geq 0 \text{ und } \max\limits_i z_i \leq a \\ 0, & \text{sonst} \end{cases}$$

(b) Wegen $\min(6, 18, 11, 20, 9) = 6 \geq 0$, $\max(6, 18, 11, 20, 9) = 20$ gilt:

$$f(6, 18, 11, 20, 9|a) = \begin{cases} \frac{1}{a^5}, & \text{für } a \geq 20 \\ 0, & \text{sonst} \end{cases}$$

$\Rightarrow f(6, 18, 11, 20, 9|a)$ maximal, falls $a \geq 20$ minimal $\Rightarrow \hat{a} = 20$