

5 Aufgaben zur Punktschätzung

Zoltán Zomotor

Versionsstand: 5. April 2016, 11:28



This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Germany License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/de/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

Aufgabe 1: Reviewfragen

- Wozu dient die Punktschätzung?
- Was bedeutet *erwartungstreu* und *asymptotisch erwartungstreu*?
- Wozu dient die Bestimmung der *Wirksamkeit* einer Schätzfunktion?
- Was bedeutet hier der Begriff *Konsistenz*?
- Was ist der Unterschied zwischen einem Maximum-Likelihood-Schätzwert und einer Maximum-Likelihood-Schätzfunktion?
- $\hat{\Theta} = f(X_1, X_2, \dots, X_n)$ sei eine Maximum-Likelihood-Schätzfunktion für den Parameter ϑ . Unter welcher Voraussetzung ist $\sin \hat{\Theta}$ eine Maximum-Likelihood-Schätzfunktion von $\sin \vartheta$?

Aufgabe 2:

Der Erwartungswert μ in der Grundgesamtheit soll durch folgende Stichprobenfunktionen geschätzt werden:

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

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

Hinweis:

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

Aufgabe 3:

Es sei x_1, \dots, x_n das Ergebnis einer einfachen Stichprobe aus einer Poisson-verteilten Grundgesamtheit mit dem unbekanntem Parameter μ .

- Bestimmen Sie die Likelihood-Funktion für dieses Stichprobenergebnis
- Zeigen Sie, dass \bar{X} die Maximum-Likelihood-Schätzfunktion für den unbekanntem Parameter μ ist.
- 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.

- Bestimmen Sie die Wahrscheinlichkeitsfunktion von X (in Abhängigkeit von N).
- Ist $2X - 1$ eine erwartungstreue Schätzfunktion für N ?

Aufgabe 5:

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

$$E(X_i) = E(X_i^2) = p$$

$$E(Y_i) = 1.5p$$

$$\text{Var}(Y_i) = 1.5p(1 - 1.5p)$$

$$E(X_i \cdot Y_i) = p^2$$

- Berechnen Sie in Abhängigkeit von p die Größen

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

- Bestimmen Sie unter den drei Schätzfunktionen

$$\hat{\theta}_1 = \frac{1}{n} \sum_{i=1}^n X_i$$

$$\hat{\theta}_2 = \frac{1}{n} \sum_{i=1}^n Y_i$$

$$\hat{\theta}_3 = \frac{2}{5n} \sum_{i=1}^n (X_i + Y_i)$$

die für p erwartungstreu, 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

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

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.

- Wie müssen a und b gewählt werden, damit Z erwartungstreu für μ ist?
- Welche unter den erwartungstreuen Schätzfunktionen Z ist die wirksamste?

Aufgabe 7:

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

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

- Bestimmen Sie α so, dass $\hat{\theta}$ erwartungstreu für μ ist. Ist die resultierende Schätzfunktion konsistent?
- Begründen Sie, ob das die wirksamste Schätzfunktion für μ ist.
- Angenommen, die Grundgesamtheit ist $N(\mu, 9)$ verteilt. Wie hoch ist dann die Wahrscheinlichkeit, dass der von $\hat{\theta}$ gelieferte Schätzwert $\hat{\theta}$ 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 λ .
- (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

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

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

- (a) Bestimmen Sie den Median t_{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}_{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 $f(z_1, z_2, \dots, z_n | d)$ für eine Stichprobenrealisation (z_1, z_2, \dots, z_n) auf.
- (b) Für $n = 5$ Entfernungsmessungen (in Armlängen) hat er sich die Stichprobenrealisation

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

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

5 Lösungen

Lösung Aufgabe 1:

Lösung Aufgabe 2:

$$\begin{aligned} E(\hat{\Theta}_n) &= E\left(\frac{1}{n} \cdot (X_1 + X_2 + \dots + X_n)\right) \\ &= \frac{1}{n} \cdot (E(X_1) + E(X_2) + \dots + E(X_n)) = \frac{1}{n} \cdot (\mu + \mu + \dots + \mu) = \frac{1}{n} \cdot (n\mu) \\ &= \mu \Rightarrow \text{erwartungstreu.} \end{aligned}$$

$$\begin{aligned} \text{Var}(\hat{\Theta}_n) &= \text{Var}\left(\frac{1}{n} \cdot (X_1 + X_2 + \dots + X_n)\right) \\ &= \frac{1}{n^2} \cdot (\text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n)) = \frac{1}{n^2} \cdot (n\sigma^2) \\ &= \frac{\sigma^2}{n} \end{aligned}$$

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

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

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

Wirksamkeit:

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

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

Lösung Aufgabe 3:

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

b)

$$\ln f(x_1, \dots, x_n | \mu) = n\bar{x} \ln \mu - n\mu - \ln \prod_{i=1}^n x_i!$$

$$\begin{aligned} \frac{d \ln f}{d \mu} &= \frac{n\bar{x}}{\mu} - n \stackrel{!}{=} 0 \\ \Rightarrow \hat{\mu} &= \bar{x} \end{aligned}$$

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

Lösung Aufgabe 4:

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

(b) Ja, weil:

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

Lösung Aufgabe 5:

(a)

$$\begin{aligned} \text{Var}(X_i) &= p(1-p) & \text{Cov}(X_i, Y_i) &= -0.5p^2 & E(X_i + Y_i) &= 2.5p \\ \text{Var}(X_i + Y_i) &= p(1-p) + 1.5p(1-1.5p) - 2 \cdot 0.5p^2 & &= 2.5p - 4.25p^2 \end{aligned}$$

(b)

$$E(\hat{\theta}_1) = E(X_i) = p \Rightarrow \hat{\theta}_1 \text{ ist erwartungstreu für } p$$

$$E(\hat{\theta}_2) = E(Y_i) = 1.5p \Rightarrow \hat{\theta}_2 \text{ ist nicht erwartungstreu für } p$$

$$E(\hat{\theta}_3) = \frac{2}{5}(E(X_i) + E(Y_i)) = p \Rightarrow \hat{\theta}_3 \text{ ist erwartungstreu für } p$$

$$\text{Var}(\hat{\theta}_1) = \frac{1}{n}p(1-p)$$

$$\text{Var}(\hat{\theta}_3) = \frac{0.4}{n}p(1-1.7p)$$

$$\text{Var}(\hat{\theta}_1) \geq \text{Var}(\hat{\theta}_3) \Leftrightarrow \frac{1}{n}p(1-p) \geq \frac{0.4}{n}p(1-1.7p) \Rightarrow p \leq 1.875$$

Wegen $p \in [0, \frac{1}{2}]$ ist $p \leq 1.875$ immer erfüllt $\Rightarrow \hat{\theta}_3$ ist wirksamer als $\hat{\theta}_1$

Lösung Aufgabe 6:

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

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

$$\frac{d\text{Var}(Z)}{da} = 3a \cdot \text{Var}(Y) - 2(1-a) \cdot \text{Var}(Y) \stackrel{!}{=} 0 \Rightarrow a = \frac{2}{5}, b = \frac{3}{5}$$

$$\frac{d^2\text{Var}(Z)}{da^2} = 5 \cdot \text{Var}(Y) > 0, \text{ also wurde ein Minimum erreicht}$$

Lösung Aufgabe 7:

$$(a) E(\hat{\theta}) = \mu \cdot (1+\alpha) \stackrel{!}{=} \mu \Leftrightarrow \alpha = 0$$

$$\text{dann: } \text{Var}(\hat{\theta}) = \frac{\text{Var}(X_i)}{n-1} \rightarrow 0 \text{ für } n \rightarrow \infty \Rightarrow \hat{\theta} \text{ ist konsistent}$$

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

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

Lösung Aufgabe 8:

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

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

$$(c) \hat{\lambda} = \bar{x} = 2$$

$$\text{diskrete Zufallsvariable: } P(X \geq 2) = 1 - F(1) = 1 - 0.406 = 0.594$$

- (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 Aufgabe 9:

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

(b)

$$f(t) = \frac{dF(t)}{dt} = \frac{be^{-b\sqrt{t}}}{2\sqrt{t}}$$

$$f(1, \dots, 4|b) = \frac{b^{10} \cdot e^{-19b}}{2^{10} \cdot \sqrt{1} \cdot \dots \cdot \sqrt{4}}$$

$$\ln f(1, \dots, 4|b) = 10 \ln b - 19b - \ln(2^{10} \cdot \sqrt{1} \cdot \dots \cdot \sqrt{4})$$

$$\frac{\partial}{\partial b} \ln f(1, \dots, 4|b) = \frac{10}{b} - 19 \stackrel{!}{=} 0$$

$$\hat{b} = \frac{10}{19} = 0.526$$

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

Lösung Aufgabe 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:

$$f(z_1, \dots, z|a) = \begin{cases} \frac{1}{a^n}, & \text{falls } \min_i z_i \geq 0 \text{ und } \max_i z_i \leq a \\ 0, & \text{sonst} \end{cases}$$

- (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 $\Rightarrow \hat{a} = 20$