Induktive Statistik

(Statistische Schätzverfahren)

Aufgabe 65: (Aufgabe 6.2 im Skript)

Gegeben sei eine Zufallsstichprobe vom Umfang n aus einer Grundgesamtheit mit Dichte

$$f(x;\lambda) = \lambda e^{-\lambda x}$$

Ermitteln Sie den ML-Schätzer für λ!

Aufgabe 66: (Aufgabe 6.4 im Skript)

Ermitteln und vergleichen Sie den MSE von S^2 und S^2_{ML}, wenn $X_1, \ldots, X_n$ eine einfache Zufallsstichprobe aus einer normalverteilten Grundgesamtheit ist.

Aufgabe 67: (Aufgabe 6.6 im Skript)

Im Rahmen eines Projekts soll die Körpergröße von Studenten untersucht werden. Hierzu wird eine Zufallsstichprobe im Umfang von $n = 10$ gezogen.
Dies liefert folgende Körpergrößen in cm:

$$176, 180, 181, 168, 177, 186, 184, 173, 182, 177.$$

(a) Schätzen Sie den unbekannten Erwartungswert der Grundgesamtheit μ!

(b) Schätzen Sie die unbekannte Varianz der Grundgesamtheit σ^2!

Aufgabe 68: (Aufgabe 6.8 im Skript)

In einer Brauerei ist eine Abfüllmenge für Bier auf eine Sollmenge von 100 l pro Faß eingestellt (Standardabweichung 0.3 l). Es ist möglich, dass es nach einiger Zeit zu Abweichungen von dieser eingestellten Sollmenge kommt, so dass die Maschine neu justiert werden muss. Bei einer zufälligen Stichprobe von 50 Fässern wurde eine durchschnittliche Abfüllmenge von 100.2 l gemessen.
Erstellen Sie ein 95%-Konfidenzintervall für μ, um zu überprüfen, ob die Maschine neu justiert werden muss!

Aufgabe 69: (Aufgabe 6.9 im Skript)

Es wurden 200 Versicherte einer großen KfZ-Haftpflichtversicherung zufällig ausgewählt und nach ihrer täglichen Fahrleistung X befragt. Die Auswertung der Befragung ergab eine durchschnittliche Fahrleistung von $\bar{x} = 25$ km pro Tag und eine Varianz von $s^2 = 128$ km^2.

Stellen Sie das 95%-Konfidenzintervall für die durchschnittliche Fahrleistung μ eines Versicherten auf!

Diese Aufgaben werden in der Übung am 12.07.2017 besprochen.

INDUKTIVE STATISTIK - ÜBUNG 8

Aufgabe 58

gegeben: $f(x) = \frac{1}{c(1+x^2)}$

$f(x)$ ist eine Dichtefunktion, wenn gilt:

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$\Rightarrow$ schon berechnet: $c = \pi$

$\Rightarrow f(x) = \frac{1}{\pi(1+x^2)}$

b) gesucht: Verteilungsfunktion $F(x)$

$$F(x) = \int_{-\infty}^{x} f(y)\,dy = \int_{-\infty}^{x} \frac{1}{\pi(1+y^2)}\,dy$$

$$= \frac{1}{\pi}\int_{-\infty}^{x} \frac{1}{(1+y^2)}\,dy = \frac{1}{\pi}\left(\arctan(y)\Big|_{-\infty}^{x}\right)$$

$$= \frac{1}{\pi}\left(\arctan(x) - \left(-\frac{\pi}{2}\right)\right) = \frac{1}{\pi}\arctan(x) + \frac{1}{2}$$

c) gesucht: Quantilfunktion $F^{-1}(y)$

$$y = \frac{1}{\pi}\arctan(x) + \frac{1}{2}$$

$$y - \frac{1}{2} = \frac{1}{\pi}\arctan(x)$$

$$\pi\left(y - \frac{1}{2}\right) = \arctan(x)$$

$$\Rightarrow \tan\left(\pi\left(y - \frac{1}{2}\right)\right) = x$$

$$\Rightarrow F^{-1}(y) = \tan\left(\pi\left(y - \frac{1}{2}\right)\right) \text{ für } y \in [0,1]$$

INDUKTIVE STATISTIK - ÜBUNG 9

Aufgabe 64

a) Zu zeigen: $E(S^2) = \sigma^2$ mit $S^2 = \frac{1}{n} \sum_{j=1}^{n} (X_j - \mu)^2$

wenn μ bekannt ist.

Beweis:

$$E(S^2) = E\left(\frac{1}{n} \sum_{j=1}^{n} (x_j - \mu)^2\right)$$

$$= E\left(\frac{1}{n} \sum_{j=1}^{n} (X_j^2 - 2X_j\mu + \mu^2)\right)$$

$$= E\left(\frac{1}{n} \sum_{j=1}^{n} X_j^2 - \frac{1}{n} \cdot 2 \sum_{j=1}^{n} x_j \mu + \frac{1}{n} \sum_{j=1}^{n} \mu^2\right)$$

$$= E\left(\frac{1}{n} \sum_{j=1}^{n} x_j^2\right) - \frac{1}{n} \mu\, 2 \sum_{j=1}^{n} E(X_j) + \frac{1}{n}\, n \cdot \mu^2$$

μ

$$= E\left(\frac{1}{n} \sum_{j=1}^{n} X_j^2\right) - \frac{1}{n} \cdot \mu \cdot 2 \cdot n \cdot \mu + \mu^2$$

$$= E\left(\frac{1}{n} \sum_{j=1}^{n} X_j^2\right) - 2\mu^2 + \mu^2$$

$$= \frac{1}{n} \sum_{j=1}^{n} E(X_j^2) - \mu^2$$

$$= \frac{1}{n} \cdot n\, E(X_j^2) - \mu^2$$

$$= E(X_j^2) - (E(X_j))^2 = Var(X_j) = \sigma^2 \quad \triangleright$$

b) Zu zeigen: $E(S^2) = \sigma^2$

aber jetzt: $S^2 = \frac{1}{n-1} \sum_{j=1}^{n} (x_j - X)^2$ und $E(X) = \mu$ ist unbekannt

Beweis:

$$E(S^2) = E\left(\frac{1}{n-1} \sum_{j=1}^{n} (X_j - \overline{X})^2\right)$$

$$= E\left(\frac{1}{n-1} \sum_{j=1}^{n} (X_j^2 - 2X_j\overline{X} + \overline{X}^2)\right)$$

$$= E\left(\frac{1}{n-1}\sum_{j=1}^{n} X_j^2 - 2\cdot\frac{1}{n-1} X_j\bar{X} + \frac{1}{n-1}\sum_{j=1}^{n}\bar{X}^2\right)$$

$$= E\left(\frac{1}{n-1}\sum_{j=1}^{n} X_j^2 - 2\cdot\frac{1}{n-1}\bar{X}\underbrace{\sum_{j=1}^{n} X_j}_{n\cdot\bar{X}} + \frac{1}{n-1}\cdot n\bar{X}^2\right)$$

$$= E\left(\frac{1}{n-1}\sum_{j=1}^{n} X_j^2 - 2\cdot\frac{1}{n-1}\bar{X}\cdot n\bar{X} + \frac{n}{n-1}\bar{X}^2\right)$$

$$= E\left(\frac{1}{n-1}\sum_{j=1}^{n} X_j^2 - 2\frac{n}{n-1}\bar{X}^2 + \frac{n}{n-1}\bar{X}^2\right)$$

$$= E\left(\frac{1}{n-1}\sum_{j=1}^{n} X_j^2 - \frac{n}{n-1}\bar{X}^2\right)$$

$$= E\left(\frac{1}{n-1}\left(\sum_{j=1}^{n} X_j^2 - n\bar{X}^2\right)\right)$$

$$= \frac{1}{n-1}\sum_{j=1}^{n} E(X_j^2) - n\cdot E(\bar{X}^2)$$

Nebenrechnung: $\sigma^2 = E(X^2) - (E(X))^2$

$$\Leftrightarrow E(X^2) = \sigma^2 + \mu^2$$

und $\sigma_{\bar{X}}^2 = E(\bar{X}^2) - (E(\bar{X}))^2$

$$\Leftrightarrow E(\bar{X}^2) = \sigma_{\bar{X}}^2 + (E(\bar{X}))^2 = \frac{\sigma^2}{n} + \mu^2$$

$$= \frac{1}{n-1}\left(\sum_{j=1}^{n}(\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)\right)$$

$$= \frac{1}{n-1}\cdot\left((n(\sigma^2+\mu^2)) - n\left(\frac{\sigma^2}{n} + \mu^2\right)\right)$$

$$= \frac{1}{n-1}\left(n\cdot\sigma^2 + n\mu^2 - (\sigma^2 + n\mu^2)\right)$$

$$= \frac{1}{n-1}\left(n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2\right) = \frac{1}{n-1}\left(n\sigma^2 - \sigma^2\right)$$

$$= \frac{1}{n-1}\left(\sigma^2(n-1)\right) = \sigma^2 \quad \triangleright$$

INDUKTIVE STATISTIK - ÜBUNG 10

Aufgabe 65

gegeben: Stichprobe vom Umfang n und Dichtefunktion : $f(x;\lambda) = \lambda e^{-\lambda x}$
gesucht: ML-Schätzer für λ!

$\triangleleft$ 1. $f(x_1, \dots, x_n | \lambda) \cdot \prod_{i=1}^{n} \lambda e^{-\lambda x_i}$

2. Bilde die log-likelihood-funktion

$$L(\lambda | x_1 \dots, x_n) = \ln\left(\prod_{i=1}^{n} \lambda e^{-\lambda x_i}\right)$$

$$= \sum_{i=1}^{n} \ln(\lambda) + \ln(e^{-\lambda x_i}) = \sum_{i=1}^{n} \ln(\lambda) + (-\lambda x_i)$$

$$= n \cdot \ln(\lambda) - \lambda \sum_{i=1}^{n} x_i$$

3. Bilde 1. Ableitung der log-likelihoodfunktion

$$\frac{dL(\lambda | x_1 \dots x_n)}{d\lambda} = \left(n \cdot \ln(\lambda) - \lambda \sum_{i=1}^{n} x_i\right)$$

$$= \frac{n}{\lambda} - \sum_{i=1}^{n} x_i$$

4. Setze die 1. Ableitung $= 0$

$$0 = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i \quad \Leftrightarrow \quad \sum_{i=1}^{n} x_i = \frac{n}{\lambda}$$

$$\Leftrightarrow \lambda = \frac{n}{\sum_{i=1}^{n} x_i} \quad \Leftrightarrow \quad \lambda = \frac{1}{\bar{x}} \quad \triangleright$$

Aufgabe 67

gegeben: $n = 10$ Studenten
$X \hat{=}$ Körpergröße
Stichprobe liegt vor

a) gesucht: Erwartungswert der Grundgesamtheit
verwende $\bar{X}$, da $E(\bar{X}) = \mu$!

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i = 178{,}4 = \hat{\mu}$$

✓

Eingabe in R: `x <- c(176, 180, 181, 168, 177, 186, 184, 173, 182, 177)`
`mean(x)`

b) gesucht: Varianz der Grundgesamtheit

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 = 28{,}711$$

Eingabe in R: `sqrt(var(x))` $= 28{,}711$
`sd(x)` $= 5{,}3583$

(oder per Hand): `sum((x-mean(x))^2)/(length(x)-1)`
$\Rightarrow = 28{,}7111$

Aufgabe 68

$X \hat{=}$ Abfüllmenge von Bierfässern
$\mu = 100$ Liter, $\sigma = 0{,}3$ Liter (L)
gegeben: Stichprobe vom Umfang $n = 50$

$n = 50 > 30$
$\Rightarrow$ Wende ZGWS (Zentraler Grenzwertsatz) von Lindeberg-Levy

$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

$\Rightarrow \bar{X} \sim N\left(100, \frac{0{,}3^2}{50}\right)$

$\Rightarrow \bar{X} \sim N(100, 0.0018)$

<u>Konfidenzintervall für μ:</u>

$P\left(\bar{X} - z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$

$\rightarrow \sigma_{\bar{x}}$

$= 1 - \alpha$ mit $1 - \alpha = 0{,}95$

mit $z_{1-\frac{\alpha}{2}} = z_{1-\frac{0{,}05}{2}} = \underbrace{z_{0{,}975}}_{R:\ qnorm(0{,}975)} = 1{,}96$

$\Rightarrow P\left(100{,}2 - 1{,}96 \cdot \sqrt{0{,}0018} \leq \mu \leq 100{,}2 + 1{,}96 \cdot \sqrt{0{,}0018}\right) = 0{,}95$

$P(100{,}1168 \leq \mu \leq 100{,}2832) = 0{,}95$

Mit Wsh. 95% überdeckt das Intervall $[100{,}1168; 100{,}2832]$ den unbekannten Parameter μ.

$\Rightarrow$ 100 Liter liegt nicht im berechneten Konfidenzintervall. Die Maschine sollte also neu justiert werden.

```r
#Übung 10
#Aufgabe 67

x <- c(176,180,181,168,177,186,184,173,182,177)

mean(x)

var(x)

sd(x)

sum( (x-mean(x))^2)/(length(x)-1)
```