Sitzung 26

Schätzen und Testen

Sitzung *Mathematik für Ingenieure C4: INF* vom 27. Juli 2020

Wigand Rathmann

Lehrstuhl für Angewandte Analysis
Department Mathematik
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

Fragen

Schätzen und Testen

Leitfragen

- Wie können aus einer Stichprobe Kenngrößen einer Verteilung geschätzt werden?
- Wie könnnen Parameter in einer Verteilung geschätzt werden?
- Entspricht die Schätzung unseren Erwarungen

Ziel dieses Themas

1. Sie erkennen den Zusammenhang zwischen beschreibender und schließender Statistik.
2. Sie können den Unterschied zwischen **Schätzer** und **Schätzung** erklären.
3. Sie können den Maximum-Likelihood-Schätzer anwenden.
4. Sie können Hypothesentests anwenden.

konfidenz/intervallschätzung: alpha

Stichprobe

Stichprobe und Grundgesamtheit

Wir haben eine Grundgesamtheit die durch eine Zufallsvariable X mit einer Verteilung beschrieben wird. bzw von einer Verteilungsfamilie

Wie können Erwartungswert und Varianz bestimmt werden?

Definition 9.6

Eine n-dimensionale ZV $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ heißt **mathematische Stichprobe**, wenn die ZV X_i $(i = 1, \ldots, n)$ untereinander unabhängig und identisch entsprechend der Grundgesamtheit X verteilt sind.

sprich der Grenzwertsatz ist aktiv

Aufgabe

Aufgrund einer stoch.unabh. und identisch verteilten Stichprobe $X_1, \ldots, X_n$ soll ein Parameter Θ einer unbekannten Verteilung P^X oder eine Kenngröße (Erwartungswert, Varianz) geschätzt werden.

Begriffe

$\hat{\Theta}(X_1, X_2, \ldots, X_n)$ **Schätzfunktion**, Schätzer

$\hat{\theta}(x_1, x_2, \ldots, x_n)$ **Schätzwert** realisierungen der Schätzwerte

Stichprobenfunktionen

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \text{arithmetisches Mittel} \qquad (1)$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \qquad \text{empirische Varianz} \qquad (2)$$

Begriffe

$\hat{\Theta}(X_1, X_2, \ldots, X_n)$ **Schätzfunktion**, Schätzer

$\hat{\theta}(x_1, x_2, \ldots, x_n)$ **Schätzwert**

Stichprobenfunktionen

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \text{arithmetisches Mittel} \tag{1}$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2 \qquad \text{empirische Varianz} \tag{2}$$

Veranschaulichung

$$\hat{\Theta}_1(X_1, X_2, \ldots, X_n) = \bar{X}(X_1, X_2, \ldots, X_n) = \bar{X}$$
$$\hat{\Theta}_2(X_1, X_2, \ldots, X_n) = \tilde{X}(X_1, X_2, \ldots, X_n) = \tilde{X}$$

median

sind **Schätzfunktionen** von $\Theta = \mathrm{E}\,X$ der Grundgesamtheit X.

$$\hat{\theta}_1(x_1, x_2, \ldots, x_n) = \bar{x}(x_1, x_2, \ldots, x_n) = \boxed{\bar{x}}$$
$$\hat{\theta}_2(x_1, x_2, \ldots, x_n) = \tilde{x}(x_1, x_2, \ldots, x_n) = \boxed{\tilde{x}}$$

realisierung be

sind **Schätzwerte** von $\Theta = \mathrm{E}\,X$ der Grundgesamtheit X.

Definition 10.3

Eine Schätzfunktion $\hat{\Theta}(X_1, X_2, \ldots, X_n)$ eines Parameters Θ heißt **erwartungstreu (unverzerrt)**, wenn der $\mathrm{E}(\hat{\Theta}) = \Theta$ ist.

Eine Schätzfunktion $\hat{\Theta}$ eines Parameters Θ heißt **asymptotisch erwartungstreu**, falls für wachsenden Stichprobenumfang der Grenzwert des Erwartungswertes von $\hat{\Theta}$ gleich dem Parameter Θ ist, d.h. wenn gilt:

$$\lim_{n \to \infty} \mathrm{E}\left(\hat{\Theta}(X_1, \ldots, X_n)\right) = \Theta.$$

Frage

Ist die Schätzfunktion

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

erwartungstreu?

Definition 10.3

Eine Schätzfunktion $\hat{\Theta}(X_1, X_2, \ldots, X_n)$ eines Parameters Θ heißt **erwartungstreu (unverzerrt)**, wenn der $E(\hat{\Theta}) = \Theta$ ist.
Eine Schätzfunktion $\hat{\Theta}$ eines Parameters Θ heißt **asymptotisch erwartungstreu**, falls für wachsenden Stichprobenumfang der Grenzwert des Erwartungswertes von $\hat{\Theta}$ gleich dem Parameter Θ ist, d.h. wenn gilt:

$$\lim_{n \to \infty} E\left(\hat{\Theta}(X_1, \ldots, X_n)\right) = \Theta.$$

Frage

Ist die Schätzfunktion

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

erwartungstreu?

E[\overline{X}] == E[X]

E[\overline{X}] = E[1/n SUM X_i] = 1/n * E[SUM X_i] = 1/n *n EX = EXAlso ist das

Ausgangspunkt

Es sei eine Stichprobe $(x_1, x_2, \ldots, x_n)$ vom Umfang n aus einer Grundgesamtheit X gegeben. Der Verteilungstyp sei bekannt, die Parameter Θ_i, $(i = 1, \ldots, m)$ seien unbekannt und sollen geschätzt werden. Es wird nur ein Parameter betrachtet.

Beispiel

Schätzen des Parameters p einer Bernoulli-p-Verteilung. Dazu wird das Experiment n-mal durchgeführt und ist $m \leqslant n$ erfolgreich.

L(x_1,.., x_n) = Prod^n_{i=1] P(X=x_i) = p^m(1-p)^(n-m)

Definition 10.9

$(x_1, x_2, \ldots, x_n)$ sei eine Stichprobe vom Umfang n aus einer Grundgesamtheit, die durch eine stetige ZV X mit Dichte $f^X(x;\Theta)$, $x \in \mathbb{R}$, Θ unbekannt, beschrieben wird. Die Funktion

$$L\left(x_1, x_2, \ldots, x_n; \Theta\right) := \prod_{i=1}^{n} f^X\left(x_i; \Theta\right) \tag{3}$$

heißt dann **Likelihood-Funktion** der Stichprobe.

wir lassen also Theta als Variable.z.b. Y~EXPL(y_1, .. y

Definition 10.10

$(x_1, x_2, \ldots, x_n)$ ist eine Stichprobe vom Umfang n aus einer Grundgesamtheit, die durch eine diskrete ZV X beschrieben wird. Die Einzelwahrscheinlichkeiten sind $P(X = x_i; \Theta)$, wobei Θ unbekannt ist. Die Funktion

$$L(x_1, x_2, \ldots, x_n; \Theta) := \prod_{i=1}^{n} P(X = x_i; \Theta) \tag{4}$$

heißt dann **Likelihood-Funktion** der Stichprobe.

Prinzip der Maximum-Likelihood-Methode

1. Die Likelihood-Funktion ist für jede Stichprobe eine Funktion in Θ.
2. Als Schätzwert $\hat{\theta}$ wird derjenige Wert gewählt, für den

$$\max_{\theta} L(x_1, x_2, \ldots, x_n; \Theta)$$

gilt. (L kann mehrere Maxima besitzen.)

Definition 10.10

$(x_1, x_2, \ldots, x_n)$ ist eine Stichprobe vom Umfang n aus einer Grundgesamtheit, die durch eine diskrete ZV X beschrieben wird. Die Einzelwahrscheinlichkeiten sind $P(X = x_i; \Theta)$, wobei Θ unbekannt ist. Die Funktion

$$L(x_1, x_2, \ldots, x_n; \Theta) := \prod_{i=1}^{n} P(X = x_i; \Theta) \tag{4}$$

heißt dann **Likelihood-Funktion** der Stichprobe.

Prinzip der Maximum-Likelihood-Methode

1. Die Likelihood-Funktion ist für jede Stichprobe eine Funktion in Θ.
2. Als Schätzwert $\hat{\theta}$ wird derjenige Wert gewählt, für den

$$\max_{\theta} L(x_1, x_2, \ldots, x_n; \Theta)$$

gilt. (L kann mehrere Maxima besitzen.)

Definition 10.10

$(x_1, x_2, \ldots, x_n)$ ist eine Stichprobe vom Umfang n aus einer Grundgesamtheit, die durch eine diskrete ZV X beschrieben wird. Die Einzelwahrscheinlichkeiten sind $P(X = x_i; \Theta)$, wobei Θ unbekannt ist. Die Funktion

$$L(x_1, x_2, \ldots, x_n; \Theta) := \prod_{i=1}^{n} P(X = x_i; \Theta) \tag{4}$$

heißt dann **Likelihood-Funktion** der Stichprobe.

Prinzip der Maximum-Likelihood-Methode

1. Die Likelihood-Funktion ist für jede Stichprobe eine Funktion in Θ.
2. Als Schätzwert $\hat{\theta}$ wird derjenige Wert gewählt, für den

$$\max_{\theta} L(x_1, x_2, \ldots, x_n; \Theta)$$

gilt. (L kann mehrere Maxima besitzen.)

wir maximieren die wahrscheinlichkeit, dass die Stichproben unter die

Differenzierbarkeit

Anwenden der notwendigen Bedingung, falls L differenzierbar ist

$$\frac{\mathrm{d}\,L}{\mathrm{d}\,\Theta} = 0 \tag{5}$$

Enthält die Likelihood-Funktion einen Exponentialausdruck, so ist es günstig $\ln L$ statt L zu verwenden und von der Gleichung

$$\frac{\mathrm{d}\ln L}{\mathrm{d}\,\Theta} = 0 \tag{6}$$

auszugehen.
Die Lösung $\hat{\theta} = \hat{\theta}\,(x_1, x_2, \ldots, x_n; \Theta)$ ist eine Realisierung (ein Punktschätzwert) der Schätzfunktion $\hat{\Theta} = \hat{\Theta}\,(x_1, x_2, \ldots, x_n; \Theta)$. $\hat{\Theta}$ heißt **Maximum-Likelihood-Schätzung** für Θ.

Differenzierbarkeit

Anwenden der notwendigen Bedingung, falls L differenzierbar ist

$$\frac{\mathrm{d}\,L}{\mathrm{d}\,\Theta} = 0 \tag{5}$$

Enthält die Likelihood-Funktion einen Exponentialausdruck, so ist es günstig $\ln L$ statt L zu verwenden und von der Gleichung

$$\frac{\mathrm{d}\ln L}{\mathrm{d}\,\Theta} = 0 \tag{6}$$

auszugehen.

Die Lösung $\hat{\theta} = \hat{\theta}\,(x_1, x_2, \ldots, x_n; \Theta)$ ist eine Realisierung (ein Punktschätzwert) der Schätzfunktion $\hat{\Theta} = \hat{\Theta}\,(x_1, x_2, \ldots, x_n; \Theta)$. $\hat{\Theta}$ heißt **Maximum-Likelihood-Schätzung** für Θ.

Differenzierbarkeit

Anwenden der notwendigen Bedingung, falls L differenzierbar ist

$$\frac{\mathrm{d}\,L}{\mathrm{d}\,\Theta} = 0 \tag{5}$$

Enthält die Likelihood-Funktion einen Exponentialausdruck, so ist es günstig $\ln L$ statt L zu verwenden und von der Gleichung

$$\frac{\mathrm{d}\ln L}{\mathrm{d}\,\Theta} = 0 \tag{6}$$

auszugehen.
Die Lösung $\hat{\theta} = \hat{\theta}\,(x_1, x_2, \ldots, x_n; \Theta)$ ist eine Realisierung (ein Punktschätzwert) der Schätzfunktion $\hat{\Theta} = \hat{\Theta}\,(x_1, x_2, \ldots, x_n; \Theta)$. $\hat{\Theta}$ heißt **Maximum-Likelihood-Schätzung** für Θ.

Differenzierbarkeit

Anwenden der notwendigen Bedingung, falls L differenzierbar ist

$$\frac{\mathrm{d}\,L}{\mathrm{d}\,\Theta} = 0 \tag{5}$$

Enthält die Likelihood-Funktion einen Exponentialausdruck, so ist es günstig $\ln L$ statt L zu verwenden und von der Gleichung

$$\frac{\mathrm{d}\ln L}{\mathrm{d}\,\Theta} = 0 \tag{6}$$

auszugehen.
Die Lösung $\hat{\theta} = \hat{\theta}\,(x_1, x_2, \ldots, x_n; \Theta)$ ist eine Realisierung (ein Punktschätzwert) der Schätzfunktion $\hat{\Theta} = \hat{\Theta}\,(x_1, x_2, \ldots, x_n; \Theta)$. $\hat{\Theta}$ heißt **Maximum-Likelihood-Schätzung** für Θ.

L(x_1, ..., x_n, p) = p^m (1-p)^(n-m)ln(L) = m*log(p) + (n-m)log(1-p)d/dp = m/p - (n-m)/(1-p) === 0->... -> ... = p

Warum Konfidenzschätzung?

Bisher wurde nur der Parameter Θ geschätzt. Daraus lässt sich aber keine Aussage über die Genauigkeit einer solchen Schätzung ableiten. Die Abweichungen vom Wert des Parameters Θ können erheblich sein, insbesondere bei kleinen Stichproben.

Idee

Gesucht wird ein Intervall $I = (G_1, G_2)$ derart, dass gilt

$$P(G_1 < \Theta < G_2) = 1 - \alpha. \tag{7}$$

Begriffe

G_1, G_2	**Konfidenzgrenzen**
(G_1, G_2)	**Konfidenzintervall**
$1 - \alpha$	**Konfidenzniveau**
α	**Irrtumswahrscheinlichkeit**

Warum Konfidenzschätzung?

Bisher wurde nur der Parameter Θ geschätzt. Daraus lässt sich aber keine Aussage über die Genauigkeit einer solchen Schätzung ableiten. Die Abweichungen vom Wert des Parameters Θ können erheblich sein, insbesondere bei kleinen Stichproben.

Idee

Gesucht wird ein Intervall $I = (G_1, G_2)$ derart, dass gilt

$$P(G_1 < \Theta < G_2) = 1 - \alpha. \tag{7}$$

Begriffe

G_1, G_2	**Konfidenzgrenzen**
(G_1, G_2)	**Konfidenzintervall**
$1 - \alpha$	**Konfidenzniveau**
α	**Irrtumswahrscheinlichkeit**

Warum Konfidenzschätzung?

Bisher wurde nur der Parameter Θ geschätzt. Daraus lässt sich aber keine Aussage über die Genauigkeit einer solchen Schätzung ableiten. Die Abweichungen vom Wert des Parameters Θ können erheblich sein, insbesondere bei kleinen Stichproben.

Idee

Gesucht wird ein Intervall $I = (G_1, G_2)$ derart, dass gilt

$$P(G_1 < \Theta < G_2) = 1 - \alpha. \tag{7}$$

Begriffe

G_1, G_2	**Konfidenzgrenzen**
(G_1, G_2)	**Konfidenzintervall**
$1 - \alpha$	**Konfidenzniveau**
α	**Irrtumswahrscheinlichkeit**

Warum Konfidenzschätzung?

Bisher wurde nur der Parameter Θ geschätzt. Daraus lässt sich aber keine Aussage über die Genauigkeit einer solchen Schätzung ableiten. Die Abweichungen vom Wert des Parameters Θ können erheblich sein, insbesondere bei kleinen Stichproben.

Idee

Gesucht wird ein Intervall $I = (G_1, G_2)$ derart, dass gilt

$$P(G_1 < \Theta < G_2) = 1 - \alpha. \tag{7}$$

Begriffe

G_1, G_2	**Konfidenzgrenzen**
(G_1, G_2)	**Konfidenzintervall**
$1 - \alpha$	**Konfidenzniveau**
α	**Irrtumswahrscheinlichkeit**

Es wird die Konfidenzschätzung für den Erwartungswert μ einer normalverteilten Grundgesamtheit X betrachtet, wobei σ^2 bekannt ist.

Die Schätzfunktion für $\Theta = \mu$ lautet $\hat{\Theta} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Es folgt

$$P(\bar{X} - \delta < \mu < \bar{X} + \delta) = P\left(|\bar{X} - \mu| < \delta\right) = 1 - \alpha. \tag{8}$$

Es wird die Konfidenzschätzung für den Erwartungswert μ einer normalverteilten Grundgesamtheit X betrachtet, wobei σ^2 bekannt ist.

Die Schätzfunktion für $\Theta = \mu$ lautet $\hat{\Theta} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Es folgt

$$P(\bar{X} - \delta < \mu < \bar{X} + \delta) = P\left(|\bar{X} - \mu| < \delta\right) = 1 - \alpha. \tag{8}$$

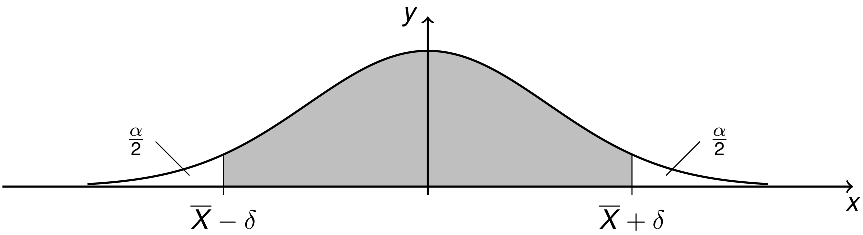
Es wird die Konfidenzschätzung für den Erwartungswert μ einer normalverteilten Grundgesamtheit X betrachtet, wobei σ^2 bekannt ist.

Die Schätzfunktion für $\Theta = \mu$ lautet $\hat{\Theta} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Es folgt

also, wenn man die Schätzung permutiert: wie gut ist der Wert tatsächlich?

$$P(\bar{X} - \delta < \mu < \bar{X} + \delta) = P\left(|\bar{X} - \mu| < \delta\right) = 1 - \alpha. \tag{8}$$



Bestimmung δ (Schätzung von μ bei bekannter Varianz σ^2)

Es ist $\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$; $Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ ist $Z \sim \mathcal{N}(0,1)$. Damit gilt

$$P\left(\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| < \frac{\delta\sqrt{n}}{\sigma}\right) = 1-\alpha \tag{9}$$

bzw.

$$P\left(|Z| < z_{1-\frac{\alpha}{2}}\right) = 1-\alpha \quad \text{mit } z_{1-\frac{\alpha}{2}} = \frac{\delta\sqrt{n}}{\sigma}. \tag{10}$$

$z_{1-\frac{\alpha}{2}}$ ist ein Perzentil der Normalverteilung und kann aus entsprechenden Tabellen abgelesen werden.

Bestimmung δ (Schätzung von μ bei bekannter Varianz σ^2)

Es ist $\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$; $Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ ist $Z \sim \mathcal{N}(0,1)$. Damit gilt

$$P\left(\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| < \frac{\delta\sqrt{n}}{\sigma}\right) = 1-\alpha \tag{9}$$

zentraler Grenzwert

bzw.

$$P\left(|Z| < z_{1-\frac{\alpha}{2}}\right) = 1-\alpha \quad \text{mit } z_{1-\frac{\alpha}{2}} = \frac{\delta\sqrt{n}}{\sigma}. \tag{10}$$

$z_{1-\frac{\alpha}{2}}$ ist ein Perzentil der Normalverteilung und kann aus entsprechenden Tabellen abgelesen werden.

Bestimmung δ (Schätzung von μ bei bekannter Varianz σ^2)

Es ist $\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$; $Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ ist $Z \sim \mathcal{N}(0,1)$. Damit gilt

$$P\left(\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| < \frac{\delta\sqrt{n}}{\sigma}\right) = 1-\alpha \tag{9}$$

bzw.

$$P\left(|Z| < z_{1-\frac{\alpha}{2}}\right) = 1-\alpha \quad \text{mit } z_{1-\frac{\alpha}{2}} = \frac{\delta\sqrt{n}}{\sigma}. \tag{10}$$

$z_{1-\frac{\alpha}{2}}$ ist ein Perzentil der Normalverteilung und kann aus entsprechenden Tabellen abgelesen werden.

Bestimmung δ (Schätzung von μ bei bekannter Varianz σ^2)

Es ist $\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$; $Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ ist $Z \sim \mathcal{N}(0,1)$. Damit gilt

$$P\left(\left|\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| < \frac{\delta\sqrt{n}}{\sigma}\right) = 1-\alpha \tag{9}$$

bzw.

$$P\left(|Z| < z_{1-\frac{\alpha}{2}}\right) = 1-\alpha \quad \text{mit } z_{1-\frac{\alpha}{2}} = \frac{\delta\sqrt{n}}{\sigma}. \tag{10}$$

$z_{1-\frac{\alpha}{2}}$ ist ein Perzentil der Normalverteilung und kann aus entsprechenden Tabellen abgelesen werden.

Selbststudium

Quellen

- Kopien Buch: Hübner, G. Stochastik. Vieweg. Kapitel 10.1-10.4
- Skript Kapitel 9 (Rückblick und Übersicht von bestimmten Summenverteilungen)
  Skript Kapitel 10.1-10.3 (Punkt- und Konfidenzschätzung)

Hinweis Buch und Skript passen an dieser Stelle nicht zusammen.

Fragen

1. Warum ist die empirische Varianz aus der beschreibenden Statistik ein „besserer Schätzer“als die Stichprobenvarianz?
2. Warum ist die Länge des Konfidenzintervalls umg. Prop zur Irrtumswahrsch
3. Was ist bei der Konfidenzschätzung einer normalverteilter Zufallsvariablen bei unbekannter Varianz zu beachten?
4. Wie ist ein Hypothesentest aufgebaut?

ende Kap. 11.1

Ihre Fragen

... stellen, Fragen haben keine Pause.

- in den Online-Sitzungen (Vorlesungen, Übungen),
- per Mail an wigand.rathmann@fau.de oder marius.yamakou@fau.de,
- im Forum `https://www.studon.fau.de/frm2897793.html`,
  Die Fragen, die bis Donnerstag gestellt wurden, werden am Freitag in der Online-Runde diskutiert.
- per Telefon (zu den Sprechzeiten sind wir auch im Büro)

Wigand Rathmann	09131/85-67129	Mi 11-12 Uhr
Marius Yamakou	09131/85-67127	Di 14-15 Uhr

Sprechstunde zur Mathematik für Ingenieure
Wann: dienstags 09:00 - 16:30 Uhr und donnerstags 09:00-17:00 Uhr, **Wo:**
`https://webconf.vc.dfn.de/ssim/` (Adobe Connect) und
`https://fau.zoom.us/j/91308761442` (Zoom)