Kapitel 2.7

Likelihood-basiertes Entscheiden

Dr. Andrea Wiencierz

Institut für Statistik, LMU München

7. Juli 2014

Wdh. datengestütztes Entscheidungsproblem

- die Komponenten eines datengestützten Entscheidungsproblems sind
  - datenfreies Entscheidungsproblem:

$$(\mathbb{A}, \Theta, l(\cdot)) \quad \text{oder} \quad (\mathbb{A}, \Theta, u(\cdot))$$

  - statistisches Modell (Informationsstruktur):

$$(\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta}) \quad \text{oder} \quad \mathcal{X} \ni X \sim P \in \mathcal{P} = \{P_\theta \,:\, \theta \in \Theta\}$$

  i.d.R. betrachten wir $X \in \mathcal{X} \subset \mathbb{R}^n$ mit $n \in \mathbb{N}$

Wdh. datengestütztes Entscheidungsproblem

- die Komponenten eines datengestützten Entscheidungsproblems sind
  - datenfreies Entscheidungsproblem:

$$(\mathbb{A}, \Theta, l(\cdot)) \quad \text{oder} \quad (\mathbb{A}, \Theta, u(\cdot))$$

  - statistisches Modell (Informationsstruktur):

$$(\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta}) \quad \text{oder} \quad \mathcal{X} \ni X \sim P \in \mathcal{P} = \{P_\theta \,:\, \theta \in \Theta\}$$

    i.d.R. betrachten wir $X \in \mathcal{X} \subset \mathbb{R}^n$ mit $n \in \mathbb{N}$

- klassisches datenbasiertes Entscheiden:

  Betrachte Entscheidungsfunktionen $d \,:\, \mathcal{X} \to \mathbb{A}$, die durch den erwarteten Verlust (R) oder Nutzen (U) über alle möglichen Beobachtungen evaluiert werden, und suche die diesbezügliche optimale Entscheidungsfunktion d^*.

Wdh. datengestütztes Entscheidungsproblem

- die Komponenten eines datengestützten Entscheidungsproblems sind
  - datenfreies Entscheidungsproblem:

$$(\mathbb{A}, \Theta, l(\cdot)) \quad \text{oder} \quad (\mathbb{A}, \Theta, u(\cdot))$$

  - statistisches Modell (Informationsstruktur):

$$(\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta}) \quad \text{oder} \quad \mathcal{X} \ni X \sim P \in \mathcal{P} = \{P_\theta \,:\, \theta \in \Theta\}$$

    i.d.R. betrachten wir $X \in \mathcal{X} \subset \mathbb{R}^n$ mit $n \in \mathbb{N}$

- klassisches datenbasiertes Entscheiden:

  Betrachte Entscheidungsfunktionen $d : \mathcal{X} \to \mathbb{A}$, die durch den erwarteten Verlust (R) oder Nutzen (U) über alle möglichen Beobachtungen evaluiert werden, und suche die diesbezügliche optimale Entscheidungsfunktion d^*.

- konditionales Entscheiden:

  Suche die für die beobachtete Stichprobe $X = x$ optimale Entscheidung a^*.

Das Likelihood-Prinzip

- gegeben ein statistisches Modell $(\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta})$ und eine beobachtete Stichprobe $X = x \in \mathcal{X}$ ist die Likelihood-Funktion L definiert als

$$L : \Theta \to \mathbb{R}_{\geq 0}, \quad \theta \mapsto L(\theta) = P_\theta(\{X = x\})$$

- betrachtet man als $\mathcal{P} = (P_\theta)_{\theta \in \Theta}$ eine Klasse stetiger Wsk.-Verteilungen, wird L häufig durch die zu P_θ zugehörige Dichte $f(\cdot; \theta)$ approximiert, d.h. dann ist $L(\theta) = f(x; \theta)$ für alle $\theta \in \Theta$

Das Likelihood-Prinzip

- gegeben ein statistisches Modell $(\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta\in\Theta})$ und eine beobachtete Stichprobe $X = x \in \mathcal{X}$ ist die Likelihood-Funktion L definiert als

$$L : \Theta \to \mathbb{R}_{\geq 0}, \quad \theta \mapsto L(\theta) = P_\theta(\{X = x\})$$

- betrachtet man als $\mathcal{P} = (P_\theta)_{\theta\in\Theta}$ eine Klasse stetiger Wsk.-Verteilungen, wird L häufig durch die zu P_θ zugehörige Dichte $f(\cdot;\theta)$ approximiert, d.h. dann ist $L(\theta) = f(x;\theta)$ für alle $\theta \in \Theta$
- das Likelihood-Prinzip:

  Seien $x, y \in \mathcal{X}$ zwei beobachtete Stichproben mit $L(\theta; x) = c\, L(\theta; y)$, für ein $c \in \mathbb{R}_{>0}$, so sollen die von x bzw. y abgeleiteten Schlüsse über θ gleich sein.
- daraus folgt:

  Die Likelihood-Funktion enthält die gesamte relevante Information über θ aus der beobachteten Stichprobe.
- für eine vertiefte Betrachtung siehe auch Casella and Berger (2002, 6.3) oder Rüger (1999, 2.3)

Beispiel 1 – Binomialverteilung

- Beobachtungen sind binomialverteilt mit Wsk. $\theta \in [0, 1]$, d.h. $X \sim \mathcal{B}(\theta, n)$
- die Wsk.-Funktion ist für jedes $x \in \mathcal{X} = \{0, \ldots, n\}$ gegeben durch

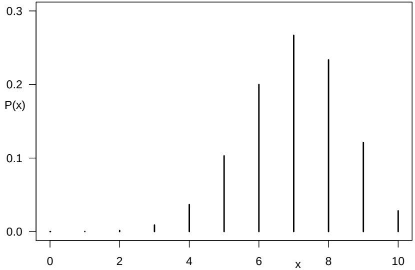
$$P_\theta(\{X = x\}) = \binom{n}{x} \theta^x \, (1 - \theta)^{n-x}$$

Beispiel 1 – Binomialverteilung

- Beobachtungen sind binomialverteilt mit Wsk. $\theta \in [0,1]$, d.h. $X \sim \mathcal{B}(\theta, n)$
- die Wsk.-Funktion ist für jedes $x \in \mathcal{X} = \{0, \ldots, n\}$ gegeben durch

$$P_\theta(\{X = x\}) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

- z.B. $\theta = 0.7$ und $n = 10$



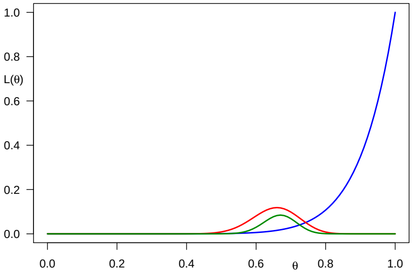
Wsk.-Masse der $\mathcal{B}(0.7, 10)$-Verteilung

Beispiel 1 – Likelihood-Funktionen

- sei nun $X \sim \mathcal{B}(0.7, n)$, betrachte 3 Realisierungen von Stichproben mit
  - $n_1 = 10$ Beobachtungen, davon 10 Erfolge
  - $n_2 = 50$ Beobachtungen, davon 33 Erfolge
  - $n_3 = 100$ Beobachtungen, davon 67 Erfolge

Beispiel 1 – Likelihood-Funktionen

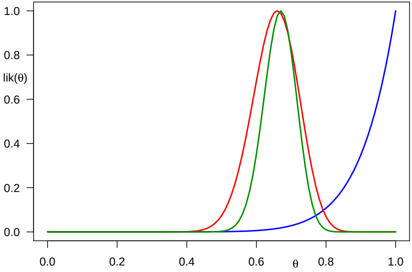
- sei nun $X \sim \mathcal{B}(0.7, n)$, betrachte 3 Realisierungen von Stichproben mit
  - $n_1 = 10$ Beobachtungen, davon 10 Erfolge
  - $n_2 = 50$ Beobachtungen, davon 33 Erfolge
  - $n_3 = 100$ Beobachtungen, davon 67 Erfolge



Likelihood-Funktionen L für $n_1 = 10$, $n_2 = 50$ und $n_3 = 100$

Beispiel 1 – Likelihood-Funktionen

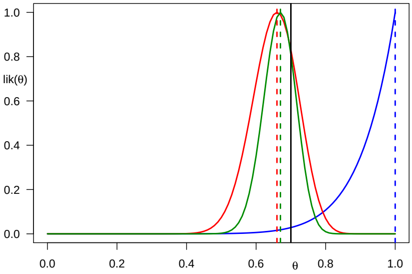
- sei nun $X \sim \mathcal{B}(0.7, n)$, betrachte 3 Realisierungen von Stichproben mit
  - $n_1 = 10$ Beobachtungen, davon 10 Erfolge
  - $n_2 = 50$ Beobachtungen, davon 33 Erfolge
  - $n_3 = 100$ Beobachtungen, davon 67 Erfolge
- normierte Likelihood $lik : \Theta \to [0,1]$, $\theta \mapsto lik(\theta) = \dfrac{L(\theta)}{\sup_{\theta' \in \Theta} L(\theta')}$



normierte Likelihood-Funktionen *lik* für $n_1 = 10$, $n_2 = 50$ und $n_3 = 100$

Beispiel 1 – Likelihood-Funktionen

- sei nun $X \sim \mathcal{B}(0.7, n)$, betrachte 3 Realisierungen von Stichproben mit
  - $n_1 = 10$ Beobachtungen, davon 10 Erfolge
  - $n_2 = 50$ Beobachtungen, davon 33 Erfolge
  - $n_3 = 100$ Beobachtungen, davon 67 Erfolge
- normierte Likelihood $lik : \Theta \to [0,1]$, $\theta \mapsto lik(\theta) = \dfrac{L(\theta)}{\sup_{\theta' \in \Theta} L(\theta')}$



normierte Likelihood-Funktionen *lik* für $n_1 = 10$, $n_2 = 50$ und $n_3 = 100$

Beispiel 2 – Normalverteilung

- Beobachtungen sind i.i.d. normalverteilt mit Erwartungswert $\mu \in \mathbb{R}$ und Varianz $\sigma^2 \in \mathbb{R}_{>0}$, d.h. $X_i \sim \mathcal{N}(\mu, \sigma^2)$ und $(\mu, \sigma^2)^{\mathrm{T}} = \theta$
- die Dichtefunktion ist für jedes $x \in \mathcal{X} = \mathbb{R}^n$ gegeben durch

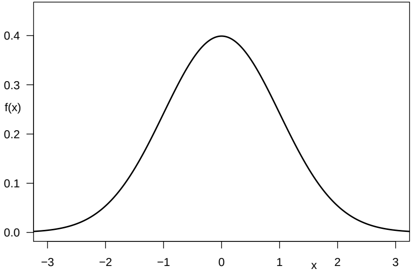
$$f(x; \theta) = \left(\frac{1}{\sqrt{2\,\pi\,\sigma^2}} \right)^n \exp\left\{ -\frac{1}{2\,\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right\}$$

Beispiel 2 – Normalverteilung

- Beobachtungen sind i.i.d. normalverteilt mit Erwartungswert $\mu \in \mathbb{R}$ und Varianz $\sigma^2 \in \mathbb{R}_{>0}$, d.h. $X_i \sim \mathcal{N}(\mu, \sigma^2)$ und $(\mu, \sigma^2)^\mathrm{T} = \theta$
- die Dichtefunktion ist für jedes $x \in \mathcal{X} = \mathbb{R}^n$ gegeben durch

$$f(x;\theta) = \left(\frac{1}{\sqrt{2\,\pi\,\sigma^2}}\right)^n \exp\left\{-\frac{1}{2\,\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}$$

- z.B. Standard-Normalverteilung, d.h. $\mu = 0$ und $\sigma^2 = 1$ für $n = 1$



Dichtefunktion der $\mathcal{N}(0,1)$-Verteilung

Beispiel 2 – Likelihood-Funktionen für μ

- sei $X = (X_1, \dots, X_n)$ mit $X_i \sim \mathcal{N}(0, 1)$ für alle $i \in \{1, \dots, n\}$
- betrachte 3 Realisierungen von Stichproben mit $n_1 = 10$, $n_2 = 50$ und $n_3 = 100$ Beobachtungen

Beispiel 2 – Likelihood-Funktionen für μ

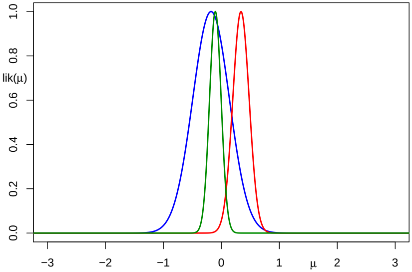
- sei $X = (X_1, \ldots, X_n)$ mit $X_i \sim \mathcal{N}(0,1)$ für alle $i \in \{1, \ldots, n\}$
- betrachte 3 Realisierungen von Stichproben mit $n_1 = 10$, $n_2 = 50$ und $n_3 = 100$ Beobachtungen
- ist $\sigma^2 = 1$ bekannt, ergibt sich die Likelihood-Funktion für μ als

$$L(\mu) = \left(\frac{1}{\sqrt{2\,\pi}}\right)^n \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)^2\right\}$$

Beispiel 2 – Likelihood-Funktionen für μ

- sei $X = (X_1, \ldots, X_n)$ mit $X_i \sim \mathcal{N}(0, 1)$ für alle $i \in \{1, \ldots, n\}$
- betrachte 3 Realisierungen von Stichproben mit $n_1 = 10$, $n_2 = 50$ und $n_3 = 100$ Beobachtungen
- ist $\sigma^2 = 1$ bekannt, ergibt sich die Likelihood-Funktion für μ als

$$L(\mu) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)^2\right\}$$



normierte Likelihood-Funktionen $lik(\mu)$ für $n_1 = 10$, $n_2 = 50$ und $n_3 = 100$

Beispiel 2 – Likelihood-Funktionen für σ^2

- sei $X = (X_1, \ldots, X_n)$ mit $X_i \sim \mathcal{N}(0, 1)$ für alle $i \in \{1, \ldots, n\}$
- betrachte 3 Realisierungen von Stichproben mit $n_1 = 10$, $n_2 = 50$ und $n_3 = 100$ Beobachtungen

Beispiel 2 – Likelihood-Funktionen für σ^2

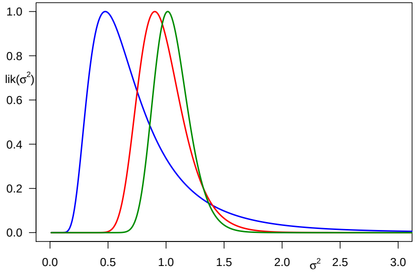
- sei $X = (X_1, \ldots, X_n)$ mit $X_i \sim \mathcal{N}(0,1)$ für alle $i \in \{1, \ldots, n\}$
- betrachte 3 Realisierungen von Stichproben mit $n_1 = 10$, $n_2 = 50$ und $n_3 = 100$ Beobachtungen
- ist $\mu = 0$ bekannt, ergibt sich die Likelihood-Funktion für σ^2 als

$$L(\sigma^2) = \left(\frac{1}{\sqrt{2\,\pi\,\sigma^2}}\right)^n \exp\left\{-\frac{1}{2\,\sigma^2}\sum_{i=1}^n x_i^2\right\}$$

Beispiel 2 – Likelihood-Funktionen für σ^2

- sei $X = (X_1, \ldots, X_n)$ mit $X_i \sim \mathcal{N}(0,1)$ für alle $i \in \{1, \ldots, n\}$
- betrachte 3 Realisierungen von Stichproben mit $n_1 = 10$, $n_2 = 50$ und $n_3 = 100$ Beobachtungen
- ist $\mu = 0$ bekannt, ergibt sich die Likelihood-Funktion für σ^2 als

$$L(\sigma^2) = \left(\frac{1}{\sqrt{2\,\pi\,\sigma^2}} \right)^n \exp\left\{ -\frac{1}{2\,\sigma^2} \sum_{i=1}^{n} x_i^2 \right\}$$



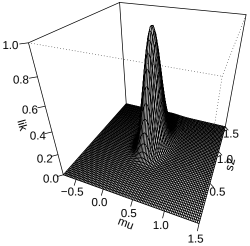
normierte Likelihood-Funktionen $lik(\sigma^2)$ für $n_1 = 10$, $n_2 = 50$ und $n_3 = 100$

Beispiel 2 – bivariate Likelihood-Funktion

- sei $X = (X_1, \dots, X_n)$ mit $X_i \sim \mathcal{N}(0,1)$ für alle $i \in \{1, \dots, n\}$
- betrachte nur die Stichprobenrealisierung mit $n_2 = 50$ Beobachtungen

Beispiel 2 – bivariate Likelihood-Funktion

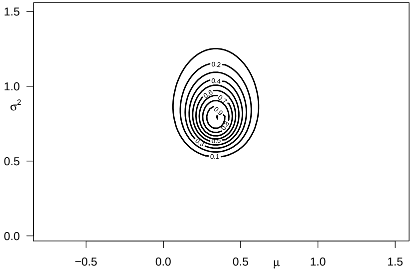
- sei $X = (X_1, \ldots, X_n)$ mit $X_i \sim \mathcal{N}(0,1)$ für alle $i \in \{1, \ldots, n\}$
- betrachte nur die Stichprobenrealisierung mit $n_2 = 50$ Beobachtungen
- bivariate Likelihood-Funktion für $\theta = (\mu, \sigma^2)^{\mathrm{T}}$



normierte Likelihood-Funktion $lik(\mu, \sigma^2)$ für $n_2 = 50$

Beispiel 2 – bivariate Likelihood-Funktion

- sei $X = (X_1, \ldots, X_n)$ mit $X_i \sim \mathcal{N}(0, 1)$ für alle $i \in \{1, \ldots, n\}$
- betrachte nur die Stichprobenrealisierung mit $n_2 = 50$ Beobachtungen
- bivariate Likelihood-Funktion für $\theta = (\mu, \sigma^2)^{\mathrm{T}}$



normierte Likelihood-Funktion $lik(\mu, \sigma^2)$ für $n_2 = 50$

Illustrationsbeispiel – Investitionsproblem

- datenfreies Problem $(\mathbb{A}, \Theta, u(\cdot))$

a_1	Investition tätigen
a_2	Investition nicht tätigen

θ_1	Besserung der Konjunktur
θ_2	Stagnation
θ_3	Konjunktur fällt

$u(a_i, \theta_j)$	θ_1	θ_2	θ_3
a_1	10 000	2 000	-15 000
a_2	1 000	1 000	0

- statistisches Modell $(\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta})$

x_1	Konjunktur wird voraussichtlich steigen
x_2	Stagnation wird erwartet
x_3	Konjunktur wird voraussichtlich fallen

$P_{\theta_j}(\{X = x_i\})$	x_1	x_2	x_3
θ_1	0.6	0.3	0.1
θ_2	0.2	0.4	0.4
θ_3	0.1	0.4	0.5

Illustrationsbeispiel – Investitionsproblem

- datenfreies Problem $(\mathbb{A}, \Theta, u(\cdot))$

a_1	Investition tätigen
a_2	Investition nicht tätigen

θ_1	Besserung der Konjunktur
θ_2	Stagnation
θ_3	Konjunktur fällt

$u(a_i, \theta_j)$	θ_1	θ_2	θ_3
a_1	10 000	2 000	-15 000
a_2	1 000	1 000	0

- statistisches Modell $(\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta})$

x_1	Konjunktur wird voraussichtlich steigen
x_2	Stagnation wird erwartet
x_3	Konjunktur wird voraussichtlich fallen

$P_{\theta_j}(\{X = x_i\})$	x_1	x_2	x_3
θ_1	0.6	0.3	0.1
θ_2	0.2	0.4	0.4
θ_3	0.1	0.4	0.5

- nehmen wir an, dass $X = x_1$ beobachtet wurde, dann nimmt die Likelihood-Funktion L folgende Werte an:

$$L(\theta_1) = P_{\theta_1}(\{X = x_1\}) = 0.6, \quad L(\theta_2) = 0.2 \quad \text{und} \quad L(\theta_3) = 0.1$$

Illustrationsbeispiel – Investitionsproblem

- datenfreies Problem $(\mathbb{A}, \Theta, u(\cdot))$

a_1	Investition tätigen
a_2	Investition nicht tätigen

θ_1	Besserung der Konjunktur
θ_2	Stagnation
θ_3	Konjunktur fällt

$u(a_i, \theta_j)$	θ_1	θ_2	θ_3
a_1	10 000	2 000	-15 000
a_2	1 000	1 000	0

- statistisches Modell $(\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta})$

x_1	Konjunktur wird voraussichtlich steigen
x_2	Stagnation wird erwartet
x_3	Konjunktur wird voraussichtlich fallen

$P_{\theta_j}(\{X = x_i\})$	x_1	x_2	x_3
θ_1	0.6	0.3	0.1
θ_2	0.2	0.4	0.4
θ_3	0.1	0.4	0.5

- nehmen wir an, dass $X = x_1$ beobachtet wurde, dann nimmt die Likelihood-Funktion L folgende Werte an:

$$L(\theta_1) = P_{\theta_1}(\{X = x_1\}) = 0.6, \quad L(\theta_2) = 0.2 \quad \text{und} \quad L(\theta_3) = 0.1$$

- normierte Likelihood: $\quad lik(\theta_1) = 1, \quad lik(\theta_2) = \dfrac{1}{3} \quad \text{und} \quad lik(\theta_3) = \dfrac{1}{6}$

Maximum-Likelihood-Aktionen

- Sei ein datengestütztes Entscheidungsproblem in Nutzenform $((\mathbb{A}, \Theta, u(\cdot)), (\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta}))$, eine Realisierung $X = x$ des Informationsbeschaffungs-experiments und die dadurch induzierte (normierte) Likelihood-Funktion *lik* gegeben.

Maximum-Likelihood-Aktionen

- Sei ein datengestütztes Entscheidungsproblem in Nutzenform $((\mathbb{A}, \Theta, u(\cdot)), (\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta\in\Theta}))$, eine Realisierung $X = x$ des Informationsbeschaffungs-experiments und die dadurch induzierte (normierte) Likelihood-Funktion *lik* gegeben.
- Das Entscheidungskriterium

$$\Phi_{ML} : \mathbb{A} \to \mathbb{R} , \quad a \mapsto \Phi_{ML}(a) = u(a, \hat{\theta})$$

mit

$$\hat{\theta} = \arg\max_{\theta\in\Theta} lik(\theta)$$

heißt Maximum-Likelihood-Entscheidungskriterium.

Maximum-Likelihood-Aktionen

- Sei ein datengestütztes Entscheidungsproblem in Nutzenform $((\mathbb{A}, \Theta, u(\cdot)), (\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta}))$, eine Realisierung $X = x$ des Informationsbeschaffungsexperiments und die dadurch induzierte (normierte) Likelihood-Funktion *lik* gegeben.
- Das Entscheidungskriterium

$$\Phi_{ML} : \mathbb{A} \to \mathbb{R} , \quad a \mapsto \Phi_{ML}(a) = u(a, \hat{\theta})$$

mit

$$\hat{\theta} = \arg\max_{\theta \in \Theta} lik(\theta)$$

heißt Maximum-Likelihood-Entscheidungskriterium.

Eine Aktion a^*, für die gilt $\Phi_{ML}(a^*) = \sup_{a \in \mathbb{A}} \Phi_{ML}(a)$, heißt Maximum-Likelihood-Aktion.

Maximum-Likelihood-Aktionen

- Sei ein datengestütztes Entscheidungsproblem in Nutzenform $((\mathbb{A}, \Theta, u(\cdot)),$ $(\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta\in\Theta}))$, eine Realisierung $X = x$ des Informationsbeschaffungs-experiments und die dadurch induzierte (normierte) Likelihood-Funktion *lik* gegeben.

- Das Entscheidungskriterium

$$\Phi_{ML} : \mathbb{A} \to \mathbb{R}\ , \quad a \mapsto \Phi_{ML}(a) = u(a, \hat{\theta})$$

mit

$$\hat{\theta} = \arg\max_{\theta\in\Theta} lik(\theta)$$

heißt Maximum-Likelihood-Entscheidungskriterium.

Eine Aktion a^*, für die gilt $\Phi_{ML}(a^*) = \sup_{a\in\mathbb{A}} \Phi_{ML}(a)$, heißt Maximum-Likelihood-Aktion.

- Liegt ein Entscheidungsproblem in Verlustform vor, werden Φ_{ML} und a^* analog definiert.

Maximum-Likelihood-Entscheidungskriterium

- Φ_{ML} ordnet jeder Handlungsalternative $a \in \mathbb{A}$ die ML-Schätzung des Nutzens bzw. Verlustes von a zu

Maximum-Likelihood-Entscheidungskriterium

- Φ_{ML} ordnet jeder Handlungsalternative $a \in \mathbb{A}$ die ML-Schätzung des Nutzens bzw. Verlustes von a zu
- Sei $\hat{\theta}$ der ML-Schätzer für $\theta \in \Theta$, dann ist für jede Funktion $g : \Theta \to \mathcal{G}$ der ML-Schätzer von $g(\theta)$ durch $g(\hat{\theta})$ gegeben.

Maximum-Likelihood-Entscheidungskriterium

- Φ_{ML} ordnet jeder Handlungsalternative $a \in \mathbb{A}$ die ML-Schätzung des Nutzens bzw. Verlustes von a zu
- Sei $\hat{\theta}$ der ML-Schätzer für $\theta \in \Theta$, dann ist für jede Funktion $g : \Theta \to \mathcal{G}$ der ML-Schätzer von $g(\theta)$ durch $g(\hat{\theta})$ gegeben.

  Die (normierte) Likelihood-Funktion lik_g für den transformierten Parameter $g(\theta)$ ist für jedes $\gamma \in \mathcal{G}$ definiert als

  $$lik_g(\gamma) = \sup_{\theta \in \Theta \,:\, g(\theta) = \gamma} lik(\theta)$$

  (Siehe z.B. Casella and Berger, 2002, 7.2.2 .)
- lik_g wird auch als Profile-Likelihood-Funktion für $g(\theta)$ bezeichnet

Maximum-Likelihood-Entscheidungskriterium

- Φ_{ML} ordnet jeder Handlungsalternative $a \in \mathbb{A}$ die ML-Schätzung des Nutzens bzw. Verlustes von a zu
- Sei $\hat{\theta}$ der ML-Schätzer für $\theta \in \Theta$, dann ist für jede Funktion $g : \Theta \to \mathcal{G}$ der ML-Schätzer von $g(\theta)$ durch $g(\hat{\theta})$ gegeben.

  Die (normierte) Likelihood-Funktion lik_g für den transformierten Parameter $g(\theta)$ ist für jedes $\gamma \in \mathcal{G}$ definiert als

  $$lik_g(\gamma) = \sup_{\theta \in \Theta : g(\theta) = \gamma} lik(\theta)$$

  (Siehe z.B. Casella and Berger, 2002, 7.2.2 .)
- lik_g wird auch als Profile-Likelihood-Funktion für $g(\theta)$ bezeichnet
- im Kontext der Entscheidungstheorie entspricht g für jede Aktion $a \in \mathbb{A}$ der aktionsspezifischen Nutzenfunktion $u_a = u(a, \cdot)$ bzw. Verlustfunktion $l_a = l(a, \cdot)$

ML-Aktion im Investitionsproblem bei $X = x_1$

- Profile-Likelihood-Funktion für u_{a_1}:

$$lik_{u_{a_1}}(v) = \sup_{\theta \in \Theta \,:\, u_{a_1}(\theta) = v} lik(\theta) \ , \quad \text{für } v \in \{-15\,000, 2\,000, 10\,000\}$$

ML-Aktion im Investitionsproblem bei $X = x_1$

- Profile-Likelihood-Funktion für u_{a_1}:

$$lik_{u_{a_1}}(v) = \sup_{\theta \in \Theta : u_{a_1}(\theta) = v} lik(\theta) \ , \quad \text{für } v \in \{-15\,000, 2\,000, 10\,000\}$$

$$lik_{u_{a_1}}(-15\,000) = \sup_{\theta \in \Theta : u_{a_1}(\theta) = -15\,000} lik(\theta) = lik(\theta_3) = \frac{1}{6},$$

ML-Aktion im Investitionsproblem bei $X = x_1$

- Profile-Likelihood-Funktion für u_{a_1}:

$$lik_{u_{a_1}}(v) = \sup_{\theta \in \Theta \,:\, u_{a_1}(\theta) = v} lik(\theta)\ , \quad \text{für } v \in \{-15\,000, 2\,000, 10\,000\}$$

$$lik_{u_{a_1}}(-15\,000) = \sup_{\theta \in \Theta \,:\, u_{a_1}(\theta) = -15\,000} lik(\theta) = lik(\theta_3) = \frac{1}{6},$$

$$lik_{u_{a_1}}(2\,000) = lik(\theta_2) = \frac{1}{3} \quad \text{und} \quad lik_{u_{a_1}}(10\,000) = lik(\theta_1) = 1$$

ML-Aktion im Investitionsproblem bei $X = x_1$

- Profile-Likelihood-Funktion für u_{a_1}:

$$lik_{u_{a_1}}(v) = \sup_{\theta \in \Theta \,:\, u_{a_1}(\theta) = v} lik(\theta) \ , \quad \text{für } v \in \{-15\,000, 2\,000, 10\,000\}$$

$$lik_{u_{a_1}}(-15\,000) = \sup_{\theta \in \Theta \,:\, u_{a_1}(\theta) = -15\,000} lik(\theta) = lik(\theta_3) = \frac{1}{6},$$

$$lik_{u_{a_1}}(2\,000) = lik(\theta_2) = \frac{1}{3} \quad \text{und} \quad lik_{u_{a_1}}(10\,000) = lik(\theta_1) = 1$$

- ML-Schätzer für u_{a_1} ist $10\,000 = \Phi_{ML}(a_1)$
- Profile-Likelihood-Funktion $lik_{u_{a_2}}$ für $v \in \{0, 1\,000\}$:

ML-Aktion im Investitionsproblem bei $X = x_1$

- Profile-Likelihood-Funktion für u_{a_1}:

$$lik_{u_{a_1}}(v) = \sup_{\theta \in \Theta \,:\, u_{a_1}(\theta) = v} lik(\theta) \ , \quad \text{für } v \in \{-15\,000, 2\,000, 10\,000\}$$

$$lik_{u_{a_1}}(-15\,000) = \sup_{\theta \in \Theta \,:\, u_{a_1}(\theta) = -15\,000} lik(\theta) = lik(\theta_3) = \frac{1}{6},$$

$$lik_{u_{a_1}}(2\,000) = lik(\theta_2) = \frac{1}{3} \quad \text{und} \quad lik_{u_{a_1}}(10\,000) = lik(\theta_1) = 1$$

- ML-Schätzer für u_{a_1} ist $10\,000 = \Phi_{ML}(a_1)$
- Profile-Likelihood-Funktion $lik_{u_{a_2}}$ für $v \in \{0, 1\,000\}$:

$$lik_{u_{a_2}}(0) = lik(\theta_3) = \frac{1}{6} \quad \text{und}$$

ML-Aktion im Investitionsproblem bei $X = x_1$

- Profile-Likelihood-Funktion für u_{a_1}:

$$lik_{u_{a_1}}(v) = \sup_{\theta \in \Theta \,:\, u_{a_1}(\theta) = v} lik(\theta) \;, \quad \text{für } v \in \{-15\,000, 2\,000, 10\,000\}$$

$$lik_{u_{a_1}}(-15\,000) = \sup_{\theta \in \Theta \,:\, u_{a_1}(\theta) = -15\,000} lik(\theta) = lik(\theta_3) = \frac{1}{6},$$

$$lik_{u_{a_1}}(2\,000) = lik(\theta_2) = \frac{1}{3} \quad \text{und} \quad lik_{u_{a_1}}(10\,000) = lik(\theta_1) = 1$$

- ML-Schätzer für u_{a_1} ist $10\,000 = \Phi_{ML}(a_1)$
- Profile-Likelihood-Funktion $lik_{u_{a_2}}$ für $v \in \{0, 1\,000\}$:

$$lik_{u_{a_2}}(0) = lik(\theta_3) = \frac{1}{6} \quad \text{und}$$

$$lik_{u_{a_2}}(1\,000) = \sup_{\theta \in \Theta \,:\, u_{a_2}(\theta) = 1\,000} lik(\theta) = \sup\{lik(\theta_1), lik(\theta_2)\} = 1$$

ML-Aktion im Investitionsproblem bei $X = x_1$

- Profile-Likelihood-Funktion für u_{a_1}:

$$lik_{u_{a_1}}(v) = \sup_{\theta \in \Theta \,:\, u_{a_1}(\theta) = v} lik(\theta)\ , \quad \text{für } v \in \{-15\,000, 2\,000, 10\,000\}$$

$$lik_{u_{a_1}}(-15\,000) = \sup_{\theta \in \Theta \,:\, u_{a_1}(\theta) = -15\,000} lik(\theta) = lik(\theta_3) = \frac{1}{6},$$

$$lik_{u_{a_1}}(2\,000) = lik(\theta_2) = \frac{1}{3} \quad \text{und} \quad lik_{u_{a_1}}(10\,000) = lik(\theta_1) = 1$$

- ML-Schätzer für u_{a_1} ist $10\,000 = \Phi_{ML}(a_1)$
- Profile-Likelihood-Funktion $lik_{u_{a_2}}$ für $v \in \{0, 1\,000\}$:

$$lik_{u_{a_2}}(0) = lik(\theta_3) = \frac{1}{6} \quad \text{und}$$

$$lik_{u_{a_2}}(1\,000) = \sup_{\theta \in \Theta \,:\, u_{a_2}(\theta) = 1\,000} lik(\theta) = \sup\{lik(\theta_1), lik(\theta_2)\} = 1$$

- ML-Schätzer für u_{a_2} ist $1\,000 = \Phi_{ML}(a_2)$

ML-Aktion im Investitionsproblem bei $X = x_1$

- Profile-Likelihood-Funktion für u_{a_1}:

$$lik_{u_{a_1}}(v) = \sup_{\theta \in \Theta \,:\, u_{a_1}(\theta) = v} lik(\theta) \ , \quad \text{für } v \in \{-15\,000, 2\,000, 10\,000\}$$

$$lik_{u_{a_1}}(-15\,000) = \sup_{\theta \in \Theta \,:\, u_{a_1}(\theta) = -15\,000} lik(\theta) = lik(\theta_3) = \frac{1}{6},$$

$$lik_{u_{a_1}}(2\,000) = lik(\theta_2) = \frac{1}{3} \quad \text{und} \quad lik_{u_{a_1}}(10\,000) = lik(\theta_1) = 1$$

- ML-Schätzer für u_{a_1} ist $10\,000 = \Phi_{ML}(a_1)$
- Profile-Likelihood-Funktion $lik_{u_{a_2}}$ für $v \in \{0, 1\,000\}$:

$$lik_{u_{a_2}}(0) = lik(\theta_3) = \frac{1}{6} \quad \text{und}$$

$$lik_{u_{a_2}}(1\,000) = \sup_{\theta \in \Theta \,:\, u_{a_2}(\theta) = 1\,000} lik(\theta) = \sup\{lik(\theta_1), lik(\theta_2)\} = 1$$

- ML-Schätzer für u_{a_2} ist $1\,000 = \Phi_{ML}(a_2)$
- da $\sup_{a \in \mathbb{A}} \Phi_{ML}(a) = \sup\{\Phi_{ML}(a_1), \Phi_{ML}(a_2)\} = \Phi_{ML}(a_1)$, ist $a^* = a_1$

Likelihood-based Region Minimax-Aktionen

- Sei ein datengestütztes Entscheidungsproblem in Nutzenform $((\mathbb{A}, \Theta, u(\cdot)), (\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta}))$, eine Realisierung $X = x$ des Informationsbeschaffungs-experiments und die dadurch induzierte (normierte) Likelihood-Funktion *lik* gegeben.

Likelihood-based Region Minimax-Aktionen

- Sei ein datengestütztes Entscheidungsproblem in Nutzenform $((\mathbb{A}, \Theta, u(\cdot)),$ $(\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta}))$, eine Realisierung $X = x$ des Informationsbeschaffungs-experiments und die dadurch induzierte (normierte) Likelihood-Funktion *lik* gegeben.
- Das Entscheidungskriterium

$$\Phi_{LRM} : \mathbb{A} \to \mathbb{R} , \quad a \mapsto \Phi_{LRM}(a) = \inf_{\theta \in \mathcal{T}_{>\beta}} u(a, \theta) ,$$

wobei

$$\mathcal{T}_{>\beta} = \{\theta \in \Theta : lik(\theta) > \beta\} , \quad \text{für ein } \beta \in (0, 1) ,$$

heißt Likelihood-based Region Minimax-Entscheidungskriterium.

Likelihood-based Region Minimax-Aktionen

- Sei ein datengestütztes Entscheidungsproblem in Nutzenform $((\mathbb{A}, \Theta, u(\cdot)), (\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta}))$, eine Realisierung $X = x$ des Informationsbeschaffungs-experiments und die dadurch induzierte (normierte) Likelihood-Funktion *lik* gegeben.
- Das Entscheidungskriterium

$$\Phi_{LRM} : \mathbb{A} \to \mathbb{R} , \quad a \mapsto \Phi_{LRM}(a) = \inf_{\theta \in \mathcal{T}_{>\beta}} u(a, \theta) ,$$

wobei

$$\mathcal{T}_{>\beta} = \{\theta \in \Theta : lik(\theta) > \beta\} , \quad \text{für ein } \beta \in (0, 1) ,$$

heißt Likelihood-based Region Minimax-Entscheidungskriterium.

Eine Aktion a^*, für die gilt $\Phi_{LRM}(a^*) = \sup_{a \in \mathbb{A}} \Phi_{LRM}(a)$, heißt Likelihood-based Region Minimax-Aktion.

Likelihood-based Region Minimax-Aktionen

- Sei ein datengestütztes Entscheidungsproblem in Nutzenform $((\mathbb{A}, \Theta, u(\cdot)), (\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta}))$, eine Realisierung $X = x$ des Informationsbeschaffungsexperiments und die dadurch induzierte (normierte) Likelihood-Funktion *lik* gegeben.
- Das Entscheidungskriterium

$$\Phi_{LRM} : \mathbb{A} \to \mathbb{R} , \quad a \mapsto \Phi_{LRM}(a) = \inf_{\theta \in \mathcal{T}_{>\beta}} u(a, \theta) ,$$

  wobei

$$\mathcal{T}_{>\beta} = \{\theta \in \Theta : lik(\theta) > \beta\} , \quad \text{für ein } \beta \in (0, 1) ,$$

  heißt Likelihood-based Region Minimax-Entscheidungskriterium.
  Eine Aktion a^*, für die gilt $\Phi_{LRM}(a^*) = \sup_{a \in \mathbb{A}} \Phi_{LRM}(a)$, heißt Likelihood-based Region Minimax-Aktion.
- Liegt ein Entscheidungsproblem in Verlustform vor, können Φ_{LRM} und a^* analog definiert werden.

Likelihood-based Region Minimax-Entscheidungskriterium

- für jedes $a \in \mathbb{A}$ ist Φ Untergrenze einer Konfidenzregion $\mathcal{U}_{a,>\beta}$ für u_a:

$$\Phi_{LRM}(a) = \inf \mathcal{U}_{a,>\beta} \quad \text{mit} \quad \mathcal{U}_{a,>\beta} = \bigcup_{\theta \in \mathcal{T}_{>\beta}} u(a,\theta) = \{v \, : \, lik_{u_a}(v) > \beta\}$$

Likelihood-based Region Minimax-Entscheidungskriterium

- für jedes $a \in \mathbb{A}$ ist Φ Untergrenze einer Konfidenzregion $\mathcal{U}_{a,>\beta}$ für u_a:

$$\Phi_{LRM}(a) = \inf \mathcal{U}_{a,>\beta} \quad \text{mit} \quad \mathcal{U}_{a,>\beta} = \bigcup_{\theta \in \mathcal{T}_{>\beta}} u(a,\theta) = \{v \,:\, lik_{u_a}(v) > \beta\}$$

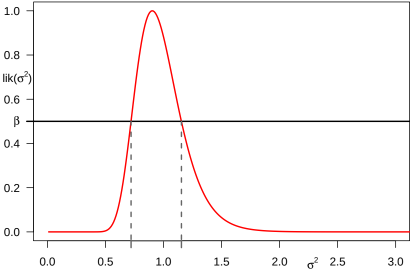
- Intervalle der Form von $\mathcal{U}_{a,>\beta}$ sind Likelihood-basierte Konfidenzregionen mit asymptotischer Überdeckungswsk. $F_{\chi^2_1}(-2\,\log(\beta))$, z.B. für $\beta = 0.5$ ca. 76%

Likelihood-based Region Minimax-Entscheidungskriterium

- für jedes $a \in \mathbb{A}$ ist Φ Untergrenze einer Konfidenzregion $\mathcal{U}_{a,>\beta}$ für u_a:

$$\Phi_{LRM}(a) = \inf \mathcal{U}_{a,>\beta} \quad \text{mit} \quad \mathcal{U}_{a,>\beta} = \bigcup_{\theta \in \mathcal{T}_{>\beta}} u(a,\theta) = \{v \,:\, lik_{u_a}(v) > \beta\}$$

- Intervalle der Form von $\mathcal{U}_{a,>\beta}$ sind Likelihood-basierte Konfidenzregionen mit asymptotischer Überdeckungswsk. $F_{\chi^2_1}(-2\log(\beta))$, z.B. für $\beta = 0.5$ ca. 76%



normierte Likelihood-Funktion $lik(\sigma^2)$ für $n_2 = 50$ mit approximativem 76% KI

LRM-Aktionen im Investitionsproblem bei $X = x_1$

- für $\beta = 0.5$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta \,:\, lik(\theta) > \beta\} = \{\theta_1\}$

LRM-Aktionen im Investitionsproblem bei $X = x_1$

- für $\beta = 0.5$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta : lik(\theta) > \beta\} = \{\theta_1\}$
- damit ist $\Phi_{LRM}(a_i) = \inf_{\theta \in \{\theta_1\}} u(a_i, \theta) = u(a_i, \hat{\theta}) = \Phi_{ML}(a_i)$ für $i \in \{1, 2\}$

LRM-Aktionen im Investitionsproblem bei $X = x_1$

- für $\beta = 0.5$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta : \mathit{lik}(\theta) > \beta\} = \{\theta_1\}$
- damit ist $\Phi_{LRM}(a_i) = \inf_{\theta \in \{\theta_1\}} u(a_i, \theta) = u(a_i, \hat{\theta}) = \Phi_{ML}(a_i)$ für $i \in \{1, 2\}$
- für $\beta = 0.26$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta : \mathit{lik}(\theta) > \beta\} = \{\theta_1, \theta_2\}$

LRM-Aktionen im Investitionsproblem bei $X = x_1$

- für $\beta = 0.5$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta : \mathit{lik}(\theta) > \beta\} = \{\theta_1\}$
- damit ist $\Phi_{LRM}(a_i) = \inf_{\theta \in \{\theta_1\}} u(a_i, \theta) = u(a_i, \hat{\theta}) = \Phi_{ML}(a_i)$ für $i \in \{1, 2\}$
- für $\beta = 0.26$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta : \mathit{lik}(\theta) > \beta\} = \{\theta_1, \theta_2\}$
- damit ist $\Phi_{LRM}(a_i) = \inf_{\theta \in \{\theta_1, \theta_2\}} u(a_i, \theta)$ für $i \in \{1, 2\}$, d.h.

$$\Phi_{LRM}(a_1) = \inf\{u(a_1, \theta_1), u(a_1, \theta_2)\} = 2\,000 \quad \text{und} \quad \Phi_{LRM}(a_2) = 1\,000$$

- also ergibt sich als LRM-Aktion $a^* = a_1$

LRM-Aktionen im Investitionsproblem bei $X = x_1$

- für $\beta = 0.5$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta : lik(\theta) > \beta\} = \{\theta_1\}$
- damit ist $\Phi_{LRM}(a_i) = \inf_{\theta \in \{\theta_1\}} u(a_i, \theta) = u(a_i, \hat{\theta}) = \Phi_{ML}(a_i)$ für $i \in \{1, 2\}$
- für $\beta = 0.26$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta : lik(\theta) > \beta\} = \{\theta_1, \theta_2\}$
- damit ist $\Phi_{LRM}(a_i) = \inf_{\theta \in \{\theta_1, \theta_2\}} u(a_i, \theta)$ für $i \in \{1, 2\}$, d.h.

$$\Phi_{LRM}(a_1) = \inf\{u(a_1, \theta_1), u(a_1, \theta_2)\} = 2\,000 \quad \text{und} \quad \Phi_{LRM}(a_2) = 1\,000$$

- also ergibt sich als LRM-Aktion $a^* = a_1$
- für $\beta = 0.15$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta : lik(\theta) > \beta\} = \{\theta_1, \theta_2, \theta_3\}$

LRM-Aktionen im Investitionsproblem bei $X = x_1$

- für $\beta = 0.5$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta : lik(\theta) > \beta\} = \{\theta_1\}$
- damit ist $\Phi_{LRM}(a_i) = \inf_{\theta \in \{\theta_1\}} u(a_i, \theta) = u(a_i, \hat{\theta}) = \Phi_{ML}(a_i)$ für $i \in \{1, 2\}$
- für $\beta = 0.26$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta : lik(\theta) > \beta\} = \{\theta_1, \theta_2\}$
- damit ist $\Phi_{LRM}(a_i) = \inf_{\theta \in \{\theta_1, \theta_2\}} u(a_i, \theta)$ für $i \in \{1, 2\}$, d.h.

$$\Phi_{LRM}(a_1) = \inf\{u(a_1, \theta_1), u(a_1, \theta_2)\} = 2\,000 \quad \text{und} \quad \Phi_{LRM}(a_2) = 1\,000$$

- also ergibt sich als LRM-Aktion $a^* = a_1$
- für $\beta = 0.15$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta : lik(\theta) > \beta\} = \{\theta_1, \theta_2, \theta_3\}$
- damit ist $\Phi_{LRM}(a_i) = \inf_{\theta \in \{\theta_1, \theta_2, \theta_3\}} u(a_i, \theta)$ für $i \in \{1, 2\}$, d.h.

$$\Phi_{LRM}(a_1) = \inf\{u(a_1, \theta_1), u(a_1, \theta_2), u(a_1, \theta_3)\} = -15\,000 \quad \text{und}$$

$$\Phi_{LRM}(a_2) = \inf\{u(a_2, \theta_1), u(a_2, \theta_2), u(a_2, \theta_3)\} = 0$$

- also ergibt sich als LRM-Aktion $a^* = a_2$

LRM-Aktionen im Investitionsproblem bei $X = x_2$

- für $X = x_2$ ist $\quad lik(\theta_1) = \frac{3}{4} \quad$ und $\quad lik(\theta_2) = lik(\theta_3) = 1$

LRM-Aktionen im Investitionsproblem bei $X = x_2$

- für $X = x_2$ ist $\quad lik(\theta_1) = \frac{3}{4} \quad$ und $\quad lik(\theta_2) = lik(\theta_3) = 1$
- für $\beta \in \{0.15, 0.26, 0.5\}$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta : lik(\theta) > \beta\} = \{\theta_1, \theta_2, \theta_3\}$

LRM-Aktionen im Investitionsproblem bei $X = x_2$

- für $X = x_2$ ist $\quad lik(\theta_1) = \frac{3}{4}$ und $lik(\theta_2) = lik(\theta_3) = 1$
- für $\beta \in \{0.15, 0.26, 0.5\}$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta : lik(\theta) > \beta\} = \{\theta_1, \theta_2, \theta_3\}$
- damit ist $\Phi_{LRM}(a_i) = \inf_{\theta \in \{\theta_1, \theta_2, \theta_3\}} u(a_i, \theta)$ für $i \in \{1, 2\}$, d.h.

$$\Phi_{LRM}(a_1) = \inf\{u(a_1, \theta_1), u(a_1, \theta_2), u(a_1, \theta_3)\} = -15\,000 \quad \text{und}$$

$$\Phi_{LRM}(a_2) = \inf\{u(a_2, \theta_1), u(a_2, \theta_2), u(a_2, \theta_3)\} = 0$$

LRM-Aktionen im Investitionsproblem bei $X = x_2$

- für $X = x_2$ ist $\quad lik(\theta_1) = \frac{3}{4} \quad$ und $\quad lik(\theta_2) = lik(\theta_3) = 1$
- für $\beta \in \{0.15, 0.26, 0.5\}$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta : lik(\theta) > \beta\} = \{\theta_1, \theta_2, \theta_3\}$
- damit ist $\Phi_{LRM}(a_i) = \inf_{\theta \in \{\theta_1, \theta_2, \theta_3\}} u(a_i, \theta)$ für $i \in \{1, 2\}$, d.h.

$$\Phi_{LRM}(a_1) = \inf\{u(a_1, \theta_1), u(a_1, \theta_2), u(a_1, \theta_3)\} = -15\,000 \quad \text{und}$$

$$\Phi_{LRM}(a_2) = \inf\{u(a_2, \theta_1), u(a_2, \theta_2), u(a_2, \theta_3)\} = 0$$

- also ergibt sich als LRM-Aktion $a^* = a_2$

LRM-Aktionen im Investitionsproblem bei $X = x_2$

- für $X = x_2$ ist $\quad lik(\theta_1) = \frac{3}{4} \quad$ und $\quad lik(\theta_2) = lik(\theta_3) = 1$
- für $\beta \in \{0.15, 0.26, 0.5\}$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta \,:\, lik(\theta) > \beta\} = \{\theta_1, \theta_2, \theta_3\}$
- damit ist $\Phi_{LRM}(a_i) = \inf_{\theta \in \{\theta_1,\theta_2,\theta_3\}} u(a_i, \theta)$ für $i \in \{1, 2\}$, d.h.

$$\Phi_{LRM}(a_1) = \inf\{u(a_1,\theta_1), u(a_1,\theta_2), u(a_1,\theta_3)\} = -15\,000 \quad \text{und}$$

$$\Phi_{LRM}(a_2) = \inf\{u(a_2,\theta_1), u(a_2,\theta_2), u(a_2,\theta_3)\} = 0$$

- also ergibt sich als LRM-Aktion $a^* = a_2$
- wählt man z.B. $\beta = 0.8$, enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta \,:\, lik(\theta) > \beta\} = \{\theta_2, \theta_3\}$

LRM-Aktionen im Investitionsproblem bei $X = x_2$

- für $X = x_2$ ist $\quad lik(\theta_1) = \dfrac{3}{4} \quad$ und $\quad lik(\theta_2) = lik(\theta_3) = 1$
- für $\beta \in \{0.15, 0.26, 0.5\}$ enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta \,:\, lik(\theta) > \beta\} = \{\theta_1, \theta_2, \theta_3\}$
- damit ist $\Phi_{LRM}(a_i) = \inf_{\theta \in \{\theta_1, \theta_2, \theta_3\}} u(a_i, \theta)$ für $i \in \{1, 2\}$, d.h.

$$\Phi_{LRM}(a_1) = \inf\{u(a_1, \theta_1), u(a_1, \theta_2), u(a_1, \theta_3)\} = -15\,000 \quad \text{und}$$

$$\Phi_{LRM}(a_2) = \inf\{u(a_2, \theta_1), u(a_2, \theta_2), u(a_2, \theta_3)\} = 0$$

- also ergibt sich als LRM-Aktion $a^* = a_2$
- wählt man z.B. $\beta = 0.8$, enthält $\mathcal{T}_{>\beta} = \{\theta \in \Theta \,:\, lik(\theta) > \beta\} = \{\theta_2, \theta_3\}$
- damit ist $\Phi_{LRM}(a_i) = \inf_{\theta \in \{\theta_2, \theta_3\}} u(a_i, \theta)$ für $i \in \{1, 2\}$, d.h.

$$\Phi_{LRM}(a_1) = \inf\{u(a_1, \theta_2), u(a_1, \theta_3)\} = -15\,000 \quad \text{und} \quad \Phi_{LRM}(a_2) = 0$$

- als LRM-Aktion ergibt sich wieder $a^* = a_2$

Lösung eines Schätzproblems mithilfe von *lik*

- statistisches Schätzproblem ist datengestütztes Entscheidungsproblem in Verlustform
$$((\mathbb{A}, \Theta, l(\cdot)), (\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta})) \ ,$$
wobei $\mathbb{A} = \Theta$ und $\mathcal{X} \subset \mathbb{R}^n$ sowie häufig
$$l \ : \ \mathbb{A} \times \Theta \to \mathbb{R}_{\geq 0} \ , \quad (a, \theta) \mapsto l(a, \theta) = (a - \theta)^2$$

Lösung eines Schätzproblems mithilfe von *lik*

- statistisches Schätzproblem ist datengestütztes Entscheidungsproblem in Verlustform
$$((\mathbb{A}, \Theta, l(\cdot)), (\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta}))\ ,$$
wobei $\mathbb{A} = \Theta$ und $\mathcal{X} \subset \mathbb{R}^n$ sowie häufig
$$l\ :\ \mathbb{A} \times \Theta \to \mathbb{R}_{\geq 0}\ , \quad (a, \theta) \mapsto l(a, \theta) = (a - \theta)^2$$
- die beobachtete Realisierung $X = x$ induziert (normierte) Likelihood-Funktion $lik\ :\ \Theta \to [0, 1]$

Lösung eines Schätzproblems mithilfe von *lik*

- statistisches Schätzproblem ist datengestütztes Entscheidungsproblem in Verlustform

$$((\mathbb{A}, \Theta, l(\cdot)), (\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta\in\Theta}))\ ,$$

wobei $\mathbb{A} = \Theta$ und $\mathcal{X} \subset \mathbb{R}^n$ sowie häufig

$$l\ :\ \mathbb{A} \times \Theta \to \mathbb{R}_{\geq 0}\ , \quad (a, \theta) \mapsto l(a, \theta) = (a - \theta)^2$$

- die beobachtete Realisierung $X = x$ induziert (normierte) Likelihood-Funktion $lik\ :\ \Theta \to [0, 1]$
- ML-Entscheidungskriterium $\Phi_{ML}(a) = l(a, \hat{\theta})$ für $\hat{\theta}$ ML-Schätzer von θ

$$\inf_{a\in\mathbb{A}} \Phi_{ML}(a) = \inf_{a\in\mathbb{A}} l(a, \hat{\theta}) = 0 \quad \text{für} \quad a = \hat{\theta}$$

Lösung eines Schätzproblems mithilfe von *lik*

- statistisches Schätzproblem ist datengestütztes Entscheidungsproblem in Verlustform
$$((\mathbb{A}, \Theta, l(\cdot)), (\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta\in\Theta}))\ ,$$
wobei $\mathbb{A} = \Theta$ und $\mathcal{X} \subset \mathbb{R}^n$ sowie häufig
$$l\ :\ \mathbb{A} \times \Theta \to \mathbb{R}_{\geq 0}\ , \quad (a, \theta) \mapsto l(a, \theta) = (a - \theta)^2$$
- die beobachtete Realisierung $X = x$ induziert (normierte) Likelihood-Funktion $lik\ :\ \Theta \to [0, 1]$
- ML-Entscheidungskriterium $\Phi_{ML}(a) = l(a, \hat{\theta})$ für $\hat{\theta}$ ML-Schätzer von θ
$$\inf_{a\in\mathbb{A}} \Phi_{ML}(a) = \inf_{a\in\mathbb{A}} l(a, \hat{\theta}) = 0 \quad \text{für} \quad a = \hat{\theta}$$
- ML-Aktion $a^* = \hat{\theta}$

Lösung eines Schätzproblems mithilfe von *lik*

- statistisches Schätzproblem ist datengestütztes Entscheidungsproblem in Verlustform
$$((\mathbb{A}, \Theta, l(\cdot)), (\mathcal{X}, \sigma(\mathcal{X}), (P_\theta)_{\theta \in \Theta}))\ ,$$
wobei $\mathbb{A} = \Theta$ und $\mathcal{X} \subset \mathbb{R}^n$ sowie häufig
$$l\ :\ \mathbb{A} \times \Theta \to \mathbb{R}_{\geq 0}\ , \quad (a, \theta) \mapsto l(a, \theta) = (a - \theta)^2$$
- die beobachtete Realisierung $X = x$ induziert (normierte) Likelihood-Funktion $lik\ :\ \Theta \to [0, 1]$
- ML-Entscheidungskriterium $\Phi_{ML}(a) = l(a, \hat{\theta})$ für $\hat{\theta}$ ML-Schätzer von θ
$$\inf_{a \in \mathbb{A}} \Phi_{ML}(a) = \inf_{a \in \mathbb{A}} l(a, \hat{\theta}) = 0 \quad \text{für} \quad a = \hat{\theta}$$
- ML-Aktion $a^* = \hat{\theta}$
- LRM-Entscheidungskriterium für ein $\beta \in (0, 1)$
$$\Phi_{LRM}(a) = \sup \mathcal{L}_{a,>\beta} \quad \text{mit} \quad \mathcal{L}_{a,>\beta} = \{k \in \mathbb{R}_{\geq 0}\ :\ lik_{l_a}(k) > \beta\}$$

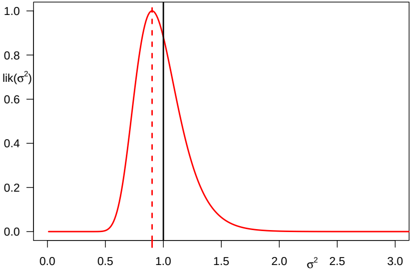
Beispiel 2 – ML-Schätzung von σ^2

- sei $X = (X_1, \dots, X_n)$ mit $X_i \sim \mathcal{N}(0, \sigma^2)$ für alle $i \in \{1, \dots, n\}$
- Schätzung von σ^2, d.h. $\mathbb{A} = \Theta = \mathbb{R}_{>0}$ mit quadratischem Verlust

Beispiel 2 – ML-Schätzung von σ^2

- sei $X = (X_1, \ldots, X_n)$ mit $X_i \sim \mathcal{N}(0, \sigma^2)$ für alle $i \in \{1, \ldots, n\}$
- Schätzung von σ^2, d.h. $\mathbb{A} = \Theta = \mathbb{R}_{>0}$ mit quadratischem Verlust
- gegeben eine Realsierung $X = x$ mit $n = n_2 = 50$ Beobachtungen ergibt sich

$$a^* = \hat{\theta} = \arg\max_{\theta \in \Theta} lik(\theta)$$



normierte Likelihood-Funktion $lik(\sigma^2)$ für $n_2 = 50$ mit ML-Schätzung

Schlussbemerkungen

- die Likelihood-basierte Entscheidungstheorie bietet eine Alternative zur Bayes-Entscheidungstheorie für konditionales, datenbasiertes Entscheiden

Schlussbemerkungen

- die Likelihood-basierte Entscheidungstheorie bietet eine Alternative zur Bayes-Entscheidungstheorie für konditionales, datenbasiertes Entscheiden
- dabei bildet die gesamte Likelihood-Funktion die Grundlage für verschiedene Likelihood-basierte Entscheidungskriterien

Schlussbemerkungen

- die Likelihood-basierte Entscheidungstheorie bietet eine Alternative zur Bayes-Entscheidungstheorie für konditionales, datenbasiertes Entscheiden
- dabei bildet die gesamte Likelihood-Funktion die Grundlage für verschiedene Likelihood-basierte Entscheidungskriterien
- Pre-Data-Eigenschaften von ML-Schätzern können im Rahmen der klassischen statistischen Entscheidungstheorie untersucht werden

Schlussbemerkungen

- die Likelihood-basierte Entscheidungstheorie bietet eine Alternative zur Bayes-Entscheidungstheorie für konditionales, datenbasiertes Entscheiden
- dabei bildet die gesamte Likelihood-Funktion die Grundlage für verschiedene Likelihood-basierte Entscheidungskriterien
- Pre-Data-Eigenschaften von ML-Schätzern können im Rahmen der klassischen statistischen Entscheidungstheorie untersucht werden
- für eine vertiefte Betrachtung der Likelihood-basierten statistischen Entscheidungstheorie siehe Cattaneo (2013, 2007)
- eine Anwendung auf das Problem der Regression mit unscharfen Beobachtungen findet sich in komprimierter Form in Cattaneo and Wiencierz (2012) und ausführlicher in Wiencierz (2013)

Literaturverzeichnis

Casella, G. and R. Berger (2002). *Statistical Inference* (2nd ed.). Duxbury.

Cattaneo, M. (2007). *Statistical Decisions Based Directly on the Likelihood Function*. Ph. D. thesis, ETH Zurich.

Cattaneo, M. (2013). Likelihood decision functions. *Electronic Journal of Statistics 7*, 2924–2946.

Cattaneo, M. and A. Wiencierz (2012). Likelihood-based Imprecise Regression. *International Journal of Approximate Reasoning 53*, 1137–1154.

Rüger, B. (1999). *Test und Schätztheorie*. Oldenbourg.

Wiencierz, A. (2013). *Regression analysis with imprecise data*. Ph. D. thesis, LMU Munich.