1.5 Datengestütztes Entscheiden – Grundbegriffe der statistischen Entscheidungstheorie

1.5.1 Ein (sehr) ausführliches Motivationsbeispiel

Stichprobeninformation, Entscheidungsfunktion und Risiko

Investitionsproblem (vgl. Bsp. 1.3.4)

Aktionen:
a_1 investieren
a_2 nicht investieren

Marketing-Investition; Erfolg hängt von zukünftiger Konjunktur ab

<u>Zustände:</u>

ϑ_1 steigende Konjunktur
ϑ_2 Stagnation
ϑ_3 fallende Konjunktur

	ϑ_1	ϑ_2	ϑ_3
a_1	10000	2000	-15000
a_2	1000	1000	0

Natürlich wird man „nicht ins Blaue“ entscheiden, sondern man wird versuchen, zusätzliche Information über die Umweltzustände zu bekommen.
Beispielweise kann man einen Konjunkturtest heranziehen.

Prognoseaussagen X mit folgenden Werten

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

Allgemein: Informationsbeschaffungsexperimente

1.5.2 Grundlegendes zur statistischen Entscheidungstheorie

Grundbegriffe

- datenfreies Problem
- Informationsstruktur
- Entscheidungsfunktion
- Risikofunktion

Def. 1.53 (Datengestütztes Entscheidungsproblem)

Ein *datengestütztes Entscheidungsproblem* in Nutzen- bzw. Verlustform ist ein Tupel

$$\begin{aligned} &\Big(\big(\mathbb{A}, \Theta, u(\cdot)\big); \big(\mathcal{X}, \sigma(\mathcal{X}), (P_\vartheta)_{\vartheta\in\Theta}\big)\Big) \\ \text{bzw.} \quad & \\ &\Big(\big(\mathbb{A}, \Theta, l(\cdot)\big); \big(\mathcal{X}, \sigma(\mathcal{X}), (P_\vartheta)_{\vartheta\in\Theta}\big)\Big) \ , \end{aligned} \tag{1.38}$$

bestehend aus den Elementen

- eines datenfreien Entscheidungsproblems $\big(\mathbb{A}, \Theta, u(\cdot)\big)$ in Nutzenform bzw. $\big(\mathbb{A}, \Theta, l(\cdot)\big)$ in Verlustform und

- eines statistischen Modells $\big(\mathcal{X}, \sigma(\mathcal{X}), (P_\vartheta)_{\vartheta\in\Theta}\big)$, dem sogenannten *Informationsbeschaffungsexperiment*, auch als *Informationsstruktur* bezeichnet.

Man beachte, dass der Zustandsraum Θ des datenfreien Problems genau der Indexmenge der Wahrscheinlichkeitmaße entspricht; das Informationsbeschaffungsexperiment soll ja (potentiell) informativ sein. Es werde implizit angenommen, dass $\sigma(\mathcal{X})$ alle Einpunktmengen $\{x\}$ mit $\{x\} \in \mathcal{X}$ enthält.

Def. und Bem. 1.54 Entscheidungsfunktionen, Auswertung datengestützter Entscheidungsprobleme

Gegeben sei ein datengestütztes Entscheidungsproblem (1.38) in Nutzenform bzw. in Verlustform.

Jede Abbildung $d : \mathcal{X} \longrightarrow \mathbb{A}$ heißt *Entscheidungsfunktion* (Strategie).

Das (formal) datenfreie Entscheidungsproblem

$$\begin{array}{ll} \big(\mathcal{D}, \Theta, \mathbf{U}\big) & \text{in Nutzenform bzw.} \\ \big(\mathcal{D}, \Theta, \mathbf{R}\big) & \text{in Verlustformform,} \end{array}$$

mit

- $\mathcal{D} \subseteq \mathbb{A}^{\mathcal{X}}$, der Menge aller Abbildungen von $\mathcal{X}$ nach $\mathbb{A}$, und
- $$\begin{aligned} \mathbf{U} : \mathcal{D} \times \Theta &\longrightarrow \mathbb{R} \\ (d, \vartheta) &\longmapsto \mathbf{U}(d, \vartheta) = \int u\big(d(x), \vartheta\big)\, dP_\vartheta(x) = \underbrace{\mathbb{E}_{P_\vartheta} u\big(d(X), \vartheta\big)} \end{aligned} \tag{1.39}$$

  bzw.

  $$\begin{aligned} \mathbf{R} : \mathcal{D} \times \Theta &\longrightarrow \mathbb{R} \\ (d, \vartheta) &\longmapsto \mathbf{R}(d, \vartheta) = \int l\big(d(x), \vartheta\big)\, dP_\vartheta(x) = \mathbb{E}_{P_\vartheta} l\big(d(X), \vartheta\big) \end{aligned} \tag{1.40}$$

heißt *zugeordnetes Auswertungsproblem* (eingeschränkt auf $\mathcal{D}$).

$\mathbf{R}(d, \vartheta)$ aus (1.40) heißt, als Funktion von ϑ für festes $d \in \mathcal{D}$ betrachtet, *Risikofunktion* der Entscheidungsfunktion d. ($\mathbf{U}(d, \vartheta)$ wird öfter als *gain* bezeichnet. In der statistischen Entscheidungstheorie, und auch hier im Folgenden, wird praktisch ausschließlich in der Verlustform gearbeitet.)

Jedes bezüglich eines Kriteriums Φ im zugeordneten Auswertungsproblem optimales Element von $\mathcal{D}$ wird als *auf* $\mathcal{D}$ *eingeschränkte* Φ*-optimale Entscheidungsfunktion* für das datengestützte Entscheidungsproblem (1.38) bezeichnet. Das heißt: Entscheidungsfunktionen werden mit Hilfe des zugeordneten Auswertungsproblems beurteilt.

Bem. 1.55

a) Da datengestützte Entscheidungsprobleme mit Hilfe des zugeordneten Auswertungsproblems ‚gelöst' werden und dieses formal die Struktur eines datenfreien Entscheidungsproblems besitzt, gelten die in Kapitel **??** gemachten Aussagen über Eigenschaften optimaler Aktionen auch für Entscheidungsfunktionen. Damit erhält man einen „Berg" an Korollaren: Zusammenhang zwischen Zulässigkeit und Bayes-Entscheidungsfunktionen, Entbehrlichkeit des Randomisierens bei Bayes-Entscheidungsfunktionen, Rolle von Äquilisatoraktionen bei der Konstruktion von Minimax-Entscheidungsproblemen: zulässige Entscheidungsfunktionen (z.B. typischerweise Bayes-Entscheidungsfunktionen) mit konstanter Risikofunktion sind Minimax-Entscheidungsfunktionen,...

b) Man kann auch entscheiden, ob sich die Berücksichtigung von Zusatzinformation, die ja normalerweise auch mit Kosten verbunden ist, lohnt: Die Differenz aus dem Kriteriumswert der optimalen Entscheidungsfunktion im datengestützten Entscheidungsproblem und dem Kriteriumswert der optimalen Aktion im datenfreien Entscheidungsproblem liefert den sogenannten *Informationswert*.

Das Auswertungsproblem

Einbettung von Test- und Schätztheorie in die statistische Entscheidungstheorie

Bem. und Bsp. 1.57 (Einbettung der Schätztheorie)

Die Aufgabe, einen Parameter aus einer entsprechenden i.i.d. Stichprobe zu schätzen, kann in die Entsscheidungstheorie eingebettet werden.

Bsp. 1.58 (Einbettung Testtheorie)

Auch die üblichen Testprobleme lassen sich in die Entscheidungstheorie einbetten.

Komplexität datengestützter Entscheidungsprobleme

Zur Randomisierung in der statistischen Entscheidungstheorie

Bem. 1.61 (Zwei Möglichkeiten zu randomisieren)

$$(\mathbb{A}, \Theta, l(\cdot)) \longrightarrow (\mathcal{D}, \Theta, \mathbf{R}(\cdot))$$

1. Die Aktionenmenge $\mathbb{A}$ besteht aus gemischten Aktionen, ist also die gemischte Erweiterung einer Menge $\mathbb{A}_0$
   Jede Entscheidungsfunktion ordnet jedem $x \in \mathcal{X}$ eine Aktion aus $\mathbb{A} = \mathcal{M}(\mathbb{A}_0)$, also eine randomisierte Aktion zu.

2. $\mathbb{A}$ besteht ausschließlich aus reinen Aktionen, im zugeordneten Auswertungsproblem $(\mathcal{D}, \Theta.\mathbf{R}(\cdot))$ wird zur gemischten Erweiterung übergegangen, d.h. es wird zwischen Entscheidungsfunktionen randomisiert. („echte randomisierte Entscheidungsfunktion", Typ B-Entscheidungsfunktion).