1.4 Randomisierte Aktionen

Def. 1.7 (randomisierte (gemischte Aktionen))

Sei $\mathbb{A}$ die Aktionenmenge eines datenfreien Entscheidungsproblems $(\mathbb{A}, \Theta, u(\cdot))$ (und $\sigma(\mathbb{A})$ eine σ-Algebra über $\mathbb{A}$, die alle Einpunktmengen $\{a\} \in \mathbb{A}$ enthält.)
Dann heißt jedes Wahrscheinlichkeitsmaß auf $(\mathbb{A}, \sigma(\mathbb{A}))$ *randomisierte (gemischte)* Aktion.
Die Menge aller randomisierten Aktionen auf $(\mathbb{A}, \sigma(\mathbb{A}))$ werde mit $\mathcal{M}^{\sigma(\mathbb{A})}(\mathbb{A})$ bezeichnet. Ist klar, welche σ-Algebra verwendet wird, so schreibt man kurz $\mathcal{M}(\mathbb{A})$.

Bem. 1.8 (Zur Semantik einer randomisierten Aktion)

Sei $\mathbb{A} = \{a_1, \ldots, a_n\}$. Dann besteht die randomisierte Aktion $\widetilde{a}(\cdot) \in \mathcal{M}(\mathbb{A})$ aus folgender Handlungsvorschrift:
Führe ein Zufallsexperiment auf $\{1, \ldots, n\}$ mit $p(\{i\}) = \widetilde{a}(\{a_i\}), \quad i = 1, \ldots, n$, durch und wähle Aktion a_i genau dann, wenn i eintritt.
D.h. es wird mit Wahrscheinlichkeit $\widetilde{a}(\{a_i\})$ die Aktion a_i gewählt.
Man schreibt oft

$$\widetilde{a} = \begin{bmatrix} a_1, & \ldots, & a_n \\ \widetilde{a}(\{a_1\}), & \ldots, & \widetilde{a}(\{a_n\}) \end{bmatrix}.$$

Wählt man $\mathcal{M}(\mathbb{A})$ als Aktionenmenge, so kann man darauf aufbauend ein eigenes Entscheidungsproblem formulieren. Dazu ist es noch nötig, den Nutzen/Verlust geeignet zu definieren (als Erwartungswert).

Def. 1.9 (gemischte Erweiterung)

Gegeben sei ein datenfreies Entscheidungsproblem $(\mathbb{A}, \Theta, u(\cdot))$ und eine geeignete σ-Algebra $\sigma(\mathbb{A})$ auf $\mathbb{A}$. Dann heißt das datenfreie Entscheidungsproblem $(\mathcal{M}^{\sigma(\mathbb{A})}(\mathbb{A}), \Theta, \widetilde{u}(\cdot))$ mit

$$\begin{aligned} \widetilde{u}(\cdot) : \mathcal{M}^{\sigma(\mathbb{A})} \times \Theta &\longrightarrow \mathbb{R} \\ (\widetilde{a}, \vartheta) \quad &\longmapsto \widetilde{u}(\widetilde{a}, \vartheta) \end{aligned}$$

und

$$\widetilde{u}(\widetilde{a}, \vartheta) := \mathbb{E}_{\widetilde{a}}[u(a, \vartheta)] = \int u(a, \vartheta) \, \mathrm{d}\widetilde{a}(a) \tag{1.6}$$

die *gemischte Erweitung* von $(\mathbb{A}, \Theta, u(\cdot))$ (bzgl. $\sigma(\mathbb{A})$).

Bem. 1.10 (zu Def.1.9)

- Die Verwendung des allgemeinen Maßintegrals erlaubt die simultane Betrachtung des stetigen und diskreten Falls sowie die Berücksichtigung gemischt stetig/diskreter Verteilungen. Insbesondere gilt:
  Ist $\widetilde{a}(\cdot)$ ein (Lebesgue-)stetiges Wahrscheinlichkeitsmaß mit Dichte $\widetilde{f}(\cdot)$, so ist
  $$\widetilde{u}(\widetilde{a}, \vartheta) = \int_{\mathbb{A}} u(a, \vartheta) \widetilde{f}(a)\, da\,. \tag{1.7}$$
  Ist $\widetilde{a}(\cdot)$ diskret mit Wahrscheinlichkeitsfunktion $\widetilde{p}(\cdot)$ auf $\mathbb{A} = \{a_1, a_2, \ldots, \}$, so ist
  $$\widetilde{u}(\widetilde{a}, \vartheta) = \sum_{i=1}^{n} u(a_i, \vartheta) \widetilde{p}(a_i)\,. \tag{1.8}$$

- Beachte Formel (1.6) ist eine definitorische Festlegung, die zwar plausibel ist, aber keineswegs logisch zwingend.

- Die gemischte Erweiterung ist also wieder ein datenfreies Entscheidungsproblem. Man braucht folglich bei allgemeinen Definitionen und Sätzen nicht zu unterscheiden, ob ein „ursprüngliches Entscheidungsproblem“ vorliegt oder die zugehörige gemischte Erweiterung.

Bem. 1.11 (reine Aktionen)

Die randomisierten Aktionen der Form $\widetilde{a}(\cdot) = \delta_{\{a\}}, \delta_{\{a\}} \in \mathbb{A}$ (Dirac-Maß = Einpunktmaß in der Menge $\{a\}$) werden als *reine Aktionen* bezeichnet und mit $a \in \mathbb{A}$ identifiziert.

Bem. 1.12 (Vom Sinn und Unsinn randomisierter Aktionen)

1.5 Entscheiden auf Datenbasis: statistische Entscheidungstheorie als Spezialfall datenfreier Entscheidungsprobleme

1.5.1 Ein (sehr) ausführliches Motivationsbeispiel

Stichprobeninformation, Entscheidungsfunktion und Risiko

Investitionsproblem (vgl. Bsp. in Abschnitt 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 erhalten. 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

Präziser: Es wird erwartet/ prognostiziert, dass ...
Allgemein: Informationsbeschaffungsexperimente

1.5.2 Grundlegendes zur statistischen Entscheidungstheorie

Grundbegriffe

- datenfreies Problem
- Informationsstruktur
- Entscheidungsfunktion
- Risikofunktion

Def. 1.13 (Datengestütztes Entscheidungsproblem)

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

$$\begin{array}{c} \Big(\big(\mathbb{A}, \Theta, u(\cdot)\big); \big(\mathcal{X}, \sigma(\mathcal{X}), (p_\vartheta)_{\vartheta\in\Theta}\big)\Big) \\ \text{bzw.} \\ \Big(\big(\mathbb{A}, \Theta, l(\cdot)\big); \big(\mathcal{X}, \sigma(\mathcal{X}), (p_\vartheta)_{\vartheta\in\Theta}\big)\Big)\ , \end{array} \tag{1.9}$$

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.

Dabei werde implizit angenommen, dass $\sigma(\mathcal{X})$ alle Einpunktmengen $\{x\}$ mit $\{x\} \in \mathcal{X}$ enthält.

Man beachte, dass der Zustandsraum Θ des datenfreien Problems genau der Indexmenge der Wahrscheinlichkeitsmaße entspricht; das Informationsbeschaffungsexperiment soll ja (potentiell) informativ sein.

Def. und Bem. 1.14 Entscheidungsfunktionen, Auswertung datengestützter Entscheidungsprobleme

Gegeben sei ein datengestütztes Entscheidungsproblem (1.9) in Nutzenform bzw. in Verlustform.

Jede (bezüglich geeigneter σ-Algebren messbare) Abbildung $d : \mathcal{X} \longrightarrow \mathbb{A}$ heißt *Entscheidungsfunktion* (Strategie).

Das (formal) datenfreie Entscheidungsproblem

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

mit

- $\emptyset \neq \mathcal{D} \subseteq \tilde{\mathbb{A}}^{\sigma(\mathcal{X})}$, der Menge aller messbaren Abbildungen von $\mathcal{X}$ nach $\mathbb{A}$, und

- $$\begin{aligned} \mathbf{U} : \mathcal{D} \times \Theta &\longrightarrow \mathbb{R} \\ (d, \vartheta) &\longmapsto \mathbf{U}(d, \vartheta) = \int u\big(d(x), \vartheta\big)\, dp_\vartheta(x) \\ &= \mathbb{E}_{p_\vartheta} \underbrace{u\big(d(X), \vartheta\big)} \end{aligned} \tag{1.10}$$

bzw.

$$\begin{aligned} \mathbf{R} : \mathcal{D} \times \Theta &\longrightarrow \mathbb{R} \\ (d, \vartheta) &\longmapsto \mathbf{R}(d, \vartheta) = \int l\big(d(x), \vartheta\big)\, dp_\vartheta(x) \\ &\qquad = \mathbb{E}_{p_\vartheta} l\big(d(X), \vartheta\big) \end{aligned} \tag{1.11}$$

heißt *zugeordnetes Auswertungsproblem* (eingeschränkt auf $\mathcal{D}$), und $(\mathbb{A}, \Theta, u(\cdot))$ bzw. $(\mathbb{A}, \Theta, l(\cdot))$ heißt dann zugehöriges *datenfreies Basisproblem.*

$\mathbf{R}(d,\vartheta)$ aus (1.11) heißt, als Funktion von ϑ für festes $d \in \mathcal{D}$ betrachtet, *Risikofunktion* der Entscheidungsfunktion d. (Die auf der Nutzenform basierende analoge Funktion $\mathbf{U}(d,\vartheta)$ wird öfter als *gain* bezeichnet. In der statistischen Entscheidungstheorie wird praktisch ausschließlich in der Verlustform gearbeitet.)

Bem. 1.15 (Auswertungsprinzip)

Entscheidungsfunktionen werden mit Hilfe des zugeordneten Auswertungsproblems beurteilt.

Bem. 1.16

a) Da datengestützte Entscheidungsprobleme mit Hilfe des zugeordneten Auswertungsproblems ‚gelöst' werden und dieses formal die Struktur eines datenfreien Entscheidungsproblems besitzt, gelten die in Kapitel **??** später gemachten Aussagen über Eigenschaften optimaler Aktionen auch für Entscheidungsfunktionen. Damit werden wir einen „Berg" an Korollaren erhalten. Wenn die Ergebnisse von Kapitel **??** allgemein meist im Kontext von [formal datenfreien] Entscheidungsproblemen formuliert werden, so gelten sie natürlich insbesondere für 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 (siehe Kapitel **??**) der optimalen Entscheidungsfunktion im datengestützten Entscheidungsproblem und dem Kriteriumswert der optimalen Aktion in dem ursprünglich zugrunde gelegten datenfreien Basisproblem liefert den sogenannten *Informationswert*.

Bsp. 1.17 (Auswertungsproblem im Investitionsproblem)

Das Auswertungsproblem Betrachten Sie das Investitionsproblem mit der in der in Kapitel 1.5 beschriebenen Informationsstruktur:

Datenfreies Basisproblem

	ϑ_1	ϑ_2	ϑ_3
a_1	10 000	2 000	–15 000
a_2	1 000	1 000	0

Informationsstruktur

	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

Notation wiederum:

$$d(i_1, i_2, \ldots, i_s) \quad \text{für} \quad d = \begin{pmatrix} x_1 & x_2 & \ldots & x_s \\ a_{i_1} & a_{i_2} & \ldots & a_{i_s} \end{pmatrix}$$

Datengestütztes Entscheidungsproblem: Auswertungsproblem

$$R(d,\vartheta) = \sum_x u(\underbrace{d(x)}_{a},\vartheta)\cdot p_\vartheta(\{x\})$$

	ϑ_1	ϑ_2	ϑ_3
$d(1,1,1)$	10 000	2 000	–15 000
$d(1,1,2)$	9 100	1 600	– 7 500
$d(1,2,1)$	7 300	1 600	– 9 000
$d(1,2,2)$	6 400	1 200	– 1 500
$d(2,1,1)$	4 600	1 800	–13 500
$d(2,1,2)$	3 700	1 400	– 6 000
$d(2,2,1)$	1 900	1 400	– 7 500
$d(2,2,2)$	1 000	1 000	0

1.5.3 Einbettung der Test- und Schätztheorie in die statistische Entscheidungstheorie

Bem. und Bsp. 1.19 (Einbettung der Schätztheorie)

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

Bem. 1.20 (Zur Einbettung von Regressionsmodellen)

Bsp. 1.21 (Einbettung Testtheorie)

Auch die üblichen Testprobleme lassen sich in die Entscheidungstheorie einbetten.

Weiteres Beispiel für Test: the lady tasting tea $\rightarrow$ spätere Übungsaufgabe
Wir werden später generell auch randomisierte Aktionen zulassen.

1.5.4 Komplexität datengestützter Entscheidungsprobleme, konditionale versus frequentisitische Sicht

1.5.5 Zur Randomisierung in der statistischen Entscheidungstheorie

Bem. 1.24 (Zwei Möglichkeiten zu randomisieren)

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

1. $\mathbb{A}$ 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“).

2. 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.