Entscheidungstheorie

T. Augustin & C. Jansen

Sommersemester 2016

AG Method(olog)ische Grundlagen der Statistik und ihre Anwendungen,
Institut für Statistik, Ludwig-Maximilians-Universität München

http://www.statistik.lmu.de/institut/ag/agmg/index.html

Sehr geehrte Leserinnen und Leser, bitte beachten Sie:

Diese Folien sind **nicht** als selbsterklärendes Skript konzipiert. Sie sollen vielmehr als Ergänzungmaterial zur Vorlesung das Mitschreiben in der Vorlesung erleichtern, indem schreibintensive Teile getippt zur Verfügung gestellt werden. Beispiele und wesentliche Zusammenhänge sollen interaktiv erarbeitet werden.

1 Einführung

1.1 Charakterisierung der Entscheidungstheorie als Theorie des rationalen Entscheidens unter Unsicherheit

1.1.1 Interdisziplinäre Bedeutung und Entwicklungsstränge

a) Rational Choice

b) Expertensysteme, entscheidungsunterstützende Systeme (Decision Support Systems)

Klir and Wierman (Uncertainty-based Information, Physika, 1998, S.1)

For three hundred years [...] uncertainty was conceived solely in terms of probability theory. This seemingly unique connection between uncertainty and probability is now challenged [... by several other] theories, which are demonstrably capable of characterizing situations under uncertainty. [...]

[...] it become clear that there are several distinct types of uncertainty. That is, it was realized that uncertainty is a multidimensional concept. [.... That] multidimensional nature of uncertainty was obscured when uncertainty was conceived solely in terms of probability theory, in which it is manifested by only one of its dimensions".

c) Statistische Entscheidungstheorie

1.1.2 Grundlegende Typen von Entscheidungsmodellen

1.2 Die Grundform eines datenfreien Entscheidungsproblems (No-data-Problem)

Def. 1.1 (Datenfreies Entscheidungsproblem)

Ein *datenfreies Entscheidungsproblem* (no-data-problem) in Nutzenform (Verlustform) ist ein Tripel $(\mathbb{A}, \Theta, u(\cdot))$ bzw. $(\mathbb{A}, \Theta, l(\cdot))$, bestehend aus

- einer Menge $\mathbb{A}$ („*Aktionenmenge*“) ,

- einer Menge Θ („*Zustandsmenge*“)

- und einer Abbildung („*Nutzenfunktion*“) (u $\hat{=}$ utility)

$$\begin{aligned} u \,:\, \mathbb{A} \times \Theta &\to \mathbb{R} \\ (a, \vartheta) &\mapsto u(a, \vartheta) \end{aligned} \tag{1.1}$$

bzw. einer Abbildung („*Verlustfunktion*“) (l $\hat{=}$ loss)

$$\begin{aligned} l \,:\, \mathbb{A} \times \Theta &\to \mathbb{R} \\ (a, \vartheta) &\mapsto l(a, \vartheta) \end{aligned} \tag{1.2}$$

Bsp. 1.2 (Das „Omelettenproblem“ von Savage)

Bem. 1.3 (Konsequenzenfunktion)

Für manche Anwendungen ist es sinnvoll, einen Schritt dazwischenzuschalten und – bei gegebenem $\mathbb{A}$ und Θ – zunächst eine *Konsequenzenfunktion*

$$\begin{aligned} c : \mathbb{A} \times \Theta &\to \mathcal{C} \\ (a, \vartheta) &\mapsto c(a, \vartheta) \end{aligned}$$

(mit $\mathcal{C}$ ist die Menge potentieller Konsequenzen) zu betrachten und darauf eine Nutzenbewertung

$$\begin{aligned} u_{\mathcal{C}} : \mathcal{C} &\to \mathbb{R} \\ c &\mapsto u_{\mathcal{C}}(c) \end{aligned}$$

bzw. eine Verlustbewertung

$$\begin{aligned} l_{\mathcal{C}} : \mathcal{C} &\to \mathbb{R} \\ l &\mapsto l_{\mathcal{C}}(c) \end{aligned}$$

festzulegen.

$u(\cdot)$ und $l(\cdot)$ ergeben sich dann durch Superposition beider Funktionen als

$$u(a, \vartheta) = u_{\mathcal{C}}(c(a, \vartheta))$$

bzw.

$$l(a, \vartheta) = l_{\mathcal{C}}(c(a, \vartheta))$$

Bem. 1.4 (zu Nutzen- und Verlustfunktionen)

- In Definition 1.1 haben Verlust- und Nutzenfunktion (als Funktionen der Gestalt $\mathbb{A} \times \Theta \to \mathbb{R}$) formal genau dieselbe Struktur. Es muss also jeweils dazugesagt werden, was vorliegen soll. (Nutzen: Je mehr, umso besser; Verlust: Je weniger, umso besser)

- Diese Uneindeutigkeit liegt nicht zuletzt daran, dass man eigentlich mit *Präferenzenordnungen* auf der Menge der Konsequenzen als grundlegende Entität arbeiten müsste. Die Nutzentheorie lehrt, wie man unter welchen Bedingungen aus Präferenzenordnungen einen kardinalen (metrischen) Nutzen konstruiert. (Hierzu kommt nur später ein kurzer Einschub.))

- Die Kardinalität das Nutzens wird (zunächst) nicht bezweifelt; man kann also mit den Nutzeneinheiten wie gewohnt rechnen. Insbesondere wird dem Folgenden eine Äquivalenz von Nutzen- und Verlustsicht unterstellt: durch Multiplizieren mit (-1) kann man dann jede Nutzenfunktion $u(\cdot)$ in eine, genau dieselbe Präferenzordnung widerspiegelnde Verlustfunktion umwandeln. Daher wird im Folgenden meist nur entweder von Nutzen- oder von Verlustfunktion gesprochen.

Bem. 1.5 (Notation im endlichen Fall; Einbettung in den $\mathbb{R}^m$)

Im Falle einer endlichen Aktionenmenge und einer endlichen Zustandsmenge wird folgende Notation verwendet:

$$\begin{aligned} \mathbb{A} &= \{a_1, ..., a_i, ..., a_n\} \\ \Theta &= \{\vartheta_1, ..., \vartheta_j, ..., \vartheta_m\} \end{aligned} \tag{1.3}$$

(also $|\mathbb{A}| = n,\ |\Theta| = m;\ i$ Laufindex in $\mathbb{A}$, j Laufindex in Θ)

Nutzenfunktion, Verlustfunktion und Konsequenzenfunktion können dann als Matrizen $(u_{ij})_{\substack{i=1,...,n \\ j=1,...,m}}$, $(l_{ij})_{\substack{i=1,...,n \\ j=1,...,m}}$, $(c_{ij})_{\substack{i=1,...,n \\ j=1,...,m}}$ dargestellt werden.

Zum Beispiel:

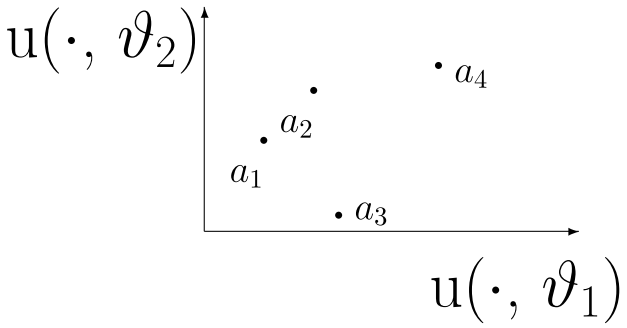
$$
\begin{array}{c|ccccc}
 & \vartheta_1 & \vartheta_2 & .. & .. & \vartheta_m \\
\hline
a_1 & u_{11} & u_{12} & .. & .. & u_{1m} \\
a_2 & u_{21} & u_{22} & .. & .. & u_{2m} \\
.. & .. & .. & .. & .. & .. \\
.. & .. & .. & u_{ij} & .. & .. \\
.. & .. & .. & .. & .. & .. \\
a_n & u_{n1} & u_{n2} & .. & .. & u_{nm}
\end{array}
\tag{1.4}
$$

Man spricht dann von Nutzen*tafel*, Verlust*tafel* oder Konsequenzen*tafel*.

Einbettung in den $\mathbb{R}^m$:
Jede Aktion $a_i \in \mathbb{A}$ kann dann mit dem zugehörigen *Nutzenvektor*

$$\vec{u}(a_i) = (u_{i1}, u_{i2}, ..., u_{ij}, ..., u_{im})^T \in \mathbb{R}^m \tag{1.5}$$

identifiziert werden.
Ist $m = 2$ oder $m = 3$, so ist ferner eine graphische Darstellung möglich:



Wird diese Sichtweise verwendet, sprechen wir von direkter Nutzen- bzw. Verlustrepräsentation der Aktionenmenge.

Bem. 1.6 (Semantik des datenfreien Entscheidungsproblems)

Abstrahiert man bei einem konkreten Entscheidungsproblem und fasst es in die Form $(\mathbb{A}, \Theta, u(\cdot))$ bzw. $(\mathbb{A}, \Theta, l(\cdot))$ laut Definition 1.1, so sind damit implizit die folgende Reihe von grundsätzlichen Annahmen verbunden, die in der jeweiligen Anwendung kritisch zu hinterfragen sind:

a) $\mathbb{A}$ sei bekannt.

b) Θ sei bekannt (Close-World-Assumption).

c) Die Ergebnisse (Konsequenzen) seien eindeutig, d.h. aus dem Zusammenspiel jedes Elements a von $\mathbb{A}$ und $\vartheta \in \Theta$ ergibt sich eine eindeutige bestimmte Konsequenz $c(a, \vartheta)$. Dabei können Konsequenzen durchaus Wahrscheinlichkeitsverteilungen sein.

d) Es läßt sich eine eindeutig bestimmte reellwertige Nutzen-/Verlustfunktion angeben, die die individuellen Präferenzen des Entscheidungsträgers vollständig widerspiegelt.

* Vgl. Bemerkung 1.4: Konstruktion von kardinalen (reellwertigen) Nutzenfunktionen aus bestimmte Bedingungen erfüllenden Präferenzordnungen $\rightarrow$ „Nutzentheorie“

* Durchgängig wird insbesondere angenommen, dass die Präferenzen eine vollständige Ordnung erlauben und in diesem Sinn eindimensional sind/gemacht werden können.

* Es gibt also keine unterschiedlichen nichtabbildbaren „Nutzendimensionen“ (z.B. eine Diskrepanz zwischen kurz- und langfristigen Nutzen). Mehrdimensionale Nutzenbewertungen sind Gegenstand der sog. multikriteriellen Entscheidungstheorie. Es gibt auch eine Reihe von Ansätzen mit unscharfer Nutzenfunktionen.[1]

[1] Für eine erste Einführung vgl. z.B. Rommelfanger & Eickemeier (2001, Kap. 2.3 und 3.1.2)

* Der Nutzen wird dabei als abstrakte Größe gesehen, in der z.B. auch Wertdispositionen mitintegriert sind.

  Aspekte wie Wertorientierung, Fairness, Reziprozität, etc. können durchaus eine Rolle spielen, sie müssen durch die Nutzenfunktion abgebildet werden bzw. abbildbar sein.

* Vorsicht: Bei monetären Konsequenzen ist im Allgemeinen der Nutzen nicht identisch mit (oder linear in) der Geldmenge. Nur wenn alle im Entscheidungsproblem betrachteten Geldmengen weit unterhalb des tatsächlichen Vermögensstandes sind, ist die Nutzenfunktion (praktisch) linear in den Geldbeträgen. Dies ist v.a. beim Berechnen von erwarteten Nutzen/Verlusten wichtig.

e) Aktionen und Zustände seien wertfrei. (Eventuelle Bewertungen müssen in die Ergebnisse und damit in die Nutzenfunktion eingebaut werden.)

f) Umwelt nicht beeinflussbar: „*Handlungsunabhängigkeit der Zustände*“
Gegebenenfalls „Implikationssschemata“ definieren (Festlegung der Umweltzustände ist keineswegs immer trivial!)

g) Typ der Unsicherheit bekannt (Typ I, Typ I’, oder Typ II, Typ II’) (bzw. später dann Verallgemeinerung)

h) Keine zusätzliche Information (außer Wahrscheinlichkeit bei Typ I, Typ I'). Informationsgewinnung über Strategien formulieren → Statistische Entscheidungstheorie

i) Einmalige Wahl der Entscheidung, keine Korrekturen

j) Keine Wiederholung der Entscheidungssituation (wiederholte Entscheidung als eine Entscheidungsstrategie formulieren)

1.3 Typische Beispiele

1.3.1 Das Ausflugsproblem nach Chernoff & Moses (1959)

(in einer Adaption von Ferschl (1975))

- Mr. Nelson möchte morgen eine Bergwanderung unternehmen, und zwar in einer Jahreszeit, in der man mit dem plötzlichen Einfallen von Schlechtwetter zu rechnen hat

- **Handlungsmöglichkeiten:**

a_1 := leichte Bekleidung mitnehmen
a_2 := leichte Bekleidung plus Regenschirm mitnehmen
a_3 := wetterfeste, warme Bekleidung plus Regenschirm mitnehmen

- Die **relevanten Fakten** bzw. **Zustände** sind

ϑ_1 := schönes Wetter am Ausflugstag
ϑ_2 := schlechtes Wetter am Ausflugstag

- Nutzentafel für Mr. Nelson:

	ϑ_1	ϑ_2
a_1	5	0
a_2	3	1
a_3	2	3

Die Bewertungen werden als Nutzen interpretiert

Potentielle datenbasierte Zusatzinfo ($\rightarrow$ später): Barometerablesung

1.3.2 Teilnahme an einer Lotterie

Gegeben sei eine Urne mit

g grünen
b blauen
r restlichen

Kugeln. Man kann

a_1 nicht spielen
a_2 zum Preis von c_g auf grün setzen
a_3 zum Preis von c_b auf blau setzen.

Nun wird eine Kugel zufällig gezogen. Ist sie grün (bzw. blau), so erhält man, wenn man auf grün (bzw. blau) gesetzt hat, w_g bzw. w_b (w: win), wobei die Beträge so klein seien, dass die Linearität des Nutzens in der Geldmenge gewährleistet sei.

$\Theta = \{\vartheta_1, \vartheta_2, \vartheta_3\}$ mit
ϑ_1 gezogene Kugel ist grün
ϑ_2 gezogene Kugel ist blau
ϑ_3 gezogene Kugel hat sonstige Farbe

Für später wichtig: Sind g,b,r bekannt, so liegt eine Unsicherheitssituation vom Typ I (Risikosituation (i.e.S.)) vor!

1.3.3 „Kuchenteilen“

Gegeben sei ein Kuchen aus 8 durchnummerierten gleich großen Stücken.

- Der Entscheidungsträger wählt ein Stück $z, z \in \{1, \ldots 7\}$, bei dem der Kuchen geteilt wird.

- Es gebe zwei Umweltzustände: ϑ_1 und ϑ_2. Tritt ϑ_1 ein, so erhält der Entscheidungsträger die Stücke 1 bis z, bei ϑ_2 hingegen die Stücke $z+1$ bis 8.

	ϑ_1	ϑ_2
1	1	7
2	2	6
3	3	5
4	4	4
5	5	3
6	6	2
7	7	1

Nutzen = Anzahl Stücke! (Nichtsättigungsannahme)

- Es wird hier noch nichts vorausgesetzt, wie die Umweltzustände eintreten. Werden sie z.B. durch einen Gegenspieler erzeugt, so erhält man ein typisches Beispiel für die Situation des strategischen Spiels, also für Typ II - Unsicherheit

1.3.4 Investitionsproblem

Ein Unternehmen steht vor der Entscheidung, eine Marketing-Investition zu tätigen. Ihr Erfolg hängt von der Konjunkturlage im nächsten Halbjahr ab.

- **Aktionen:**

$a_1 :=$ Investition tätigen

$a_2 :=$ Investition nicht tätigen

- **Zustände:**

$\vartheta_1 :=$ Steigende Konjunktur

$\vartheta_2 :=$ Stagnation

$\vartheta_3 :=$ Konjunktur fällt

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

1.3.5 Aktienkauf (natürlich stark vereinfacht)

1.3.6 Produktionsplanung unter Szenarien

- q Güter

- $\mathbb{A} \subseteq (\mathbb{R}_0^+)^q$ beschrieben durch Kapazitätsbeschränkungen

- $$a = \begin{pmatrix} a[1] \\ \vdots \\ a[q] \end{pmatrix}$$

  mit $a[\ell]$: Produktion von $a[\ell]$ Einheiten von Gut ℓ, $\ell = 1, \ldots, q$

- Umweltzustand ϑ: Szenario (mögliche Entwicklung des Marktes)

- Umsätze unter veschiedenen Szenarien ergeben Nutzenfunktion:

$$\begin{aligned} u : \mathbb{A} \times \Theta &\to \mathbb{R} \\ (a, \vartheta) &\mapsto u(a, \vartheta) \end{aligned}$$

- Einfache Form: Stückpreisbasiert; unter Nichtsättigungsannahme; Vektor $c(\vartheta)$ mit $c[\ell](\vartheta)$: Stückpreis für Gut ℓ bei Szenario ϑ, $\ell = 1, \ldots, q$, dann

$$u(a, \vartheta) = a'c(\vartheta) = \sum_{\ell=1}^{q} a[\ell] \cdot c[\ell](\vartheta)$$

bzw. bei Interpretation von $c(\vartheta)$ als Kosten:

$$\mathrm{l}(a, \vartheta) = \sum_{\ell=1}^{q} a[\ell] \cdot c[\ell](\vartheta)$$

1.3.7 Einbettung stat. Tests in die Entscheidungstheorie I: der datenfreie Kern eines Testproblems

1.3.8 Einbettung der Parameterschätzung in die Entscheidungstheorie: der datenfreie Kern eines Schätzproblems

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

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

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)

<u>Aktionen:</u>
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

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

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

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

mit

- $\emptyset \neq \mathcal{D} \subseteq \tilde{\mathbb{A}}^{\sigma(\mathcal{X})}$, der Menge aller messbaren Abbildungen von $\mathcal{X}$ nach $\mathbb{A}$, und
- $$\begin{aligned} \mathbf{U} : \mathcal{D} \times \Theta &\longrightarrow \mathbb{R} \\ (d, \vartheta) &\longmapsto \mathbf{U}(d, \vartheta) = \int u\big(d(x), \vartheta\big)\, dp_\vartheta(x) \\ &\qquad\quad = \mathbb{E}_{p_\vartheta} \underbrace{u\big(d(X), \vartheta\big)} \end{aligned} \tag{1.10}$$

  bzw.

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

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 den späteren Kapitel gemachten Aussagen über Eigenschaften optimaler Aktionen auch für Entscheidungsfunktionen. Damit werden wir einen „Berg" an Korollaren erhalten. Wenn die Ergebnisse der späteren Kapitel allgemein meist im Kontext von [formal datenfreien] Entscheidungsproblemen formuliert werden, so gelten sie natürlich insbesondere auch für datengestützte Entscheidungsproblem.

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 2) 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.

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.