Zeitschrift: Elemente der Mathematik
Herausgeber: Schweizerische Mathematische Gesellschaft
Band: 40 (1985)
Heft: 5

Artikel: Statistische Interferenz und strategisches Spiel
Autor: Loeffel, Hans
DOI: https://doi.org/10.5169/seals-38837

c) Da in der Transformationsmatrix X die Hauptvektorketten in der umgekehrten Reihenfolge auftreten, erhalten wir insgesamt folgende Ähnlichkeitstransformation $X^{-1}AX = J$:

$$\begin{pmatrix} 1 & 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & -1 & -2 & 3 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -1 & 1 & 0 \\ 1 & -1 & 2 & -1 & 2 & 0 \\ 1 & -1 & 1 & -1 & 2 & 0 \\ -1 & 0 & -4 & 5 & -4 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ -1 & 1 & 1 & -1 & 3 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Peter Lesky, Math. Institut A,
Universität Stuttgart

LITERATURVERZEICHNIS

1 A. I. Kostrikin: Introduction to Algebra. Springer-Verlag, Berlin, Heidelberg, New York 1982.
2 W. Schmeidler: Vorträge über Determinanten und Matrizen mit Anwendungen in Physik und Technik. Akademie-Verlag, Berlin 1949.
3 R. Zurmühl und S. Falk: Matrizen und ihre Anwendungen, 5. Aufl., Teil 1. Springer-Verlag, Berlin, Heidelberg, New York 1984.

 0013-6018/85/060105-05$1.50 + 0.20/0

Statistische Inferenz und strategisches Spiel

1. Einleitung

Der Entscheidungsträger (z. B. der Statistiker, der Unternehmer, der Staat usw.) sieht sich oft in der Lage, aus einer Menge $A = \{a_i\}$ [1]) von *Aktionen* oder *Entscheidungen* eine in einem gewissen Sinne *optimale* auszuwählen, und zwar angesichts einer *unsichern* Umwelt. Von dieser nehmen wir an, dass sie die relevanten Zustände $z_j \in Z$ annehmen kann. Wählt der Entscheidungsträger eine Aktion a_i im Zustand z_j, so resultiert ein Ergebnis $e_{ij} \in E$. Mit Hilfe der *Nutzenaxiomatik* nach J. v. Neumann oder Luce-Raiffa [4], S. 105ff.,

1) Wir beschränken uns auf endliche Mengen.

lässt sich der Ergebnisraum E quantifizieren, d.h. jedem Ergebnis e_{ij} wird ein Nutzen $u_{ij} \in \mathbf{R}$ oder Verlust $v_{ij} \in \mathbf{R}$ zugeordnet.
Die modellhafte Beschreibung von realen Entscheidungssituationen kann durch das sog. *Grundmodell* der Entscheidungstheorie erfolgen, das im Rahmen der statistischen Problemstellung im nachfolgenden Kapitel vorgestellt wird. Die *Entscheidungstheorie bei Unsicherheit* ist wesenhaft mit dem Zufall und dadurch mit der Wahrscheinlichkeitsrechnung und Statistik verbunden.
Angesichts einer unsichern Umwelt oder Realität ist man geneigt, diese – im Sinne einer pessimistischen Grundhaltung – als konkurrierenden Gegenspieler zu interpretieren. So gelangen wir zur Theorie der strategischen Spiele (kurz Spieltheorie genannt), die von J. v. Neumann und O. Morgenstern im klassischen Werk [1] niedergelegt wurde.

2. Die klassische Statistik im Lichte der Entscheidungstheorie

In der klassischen *induktiven* oder *inferentiellen Statistik*[2]) stehen Begriffe wie Erwartungstreue, Effizienz, Fehler 1. und 2. Art und Macht eines Tests im Mittelpunkt. Hingegen vermisst man eine ökonomische oder finanzielle *Bewertung* richtiger oder falscher Entscheide sowie den Einbezug von gewissen *Vorinformationen* bezüglich der zu schätzenden Parameter.
Epochemachend kann der universale Ansatz von A. Wald bezeichnet werden, der um 1950, ausgehend von der *Sequentialanalyse*[3]), sämtliche Probleme der statistischen Inferenz in den Rahmen der Entscheidungstheorie bei Unsicherheit eingebaut hat [2]. Im Mittelpunkt dieser Theorie steht der Begriff der *statistischen Entscheidungsfunktion* oder *Strategie.* Es ist eines unserer Ziele, diese Theorie an einem abgegrenzten Problemkreis exemplarisch darzustellen (Kapitel 5) und die verschiedenen Lösungsansätze aufzuzeigen (Kapitel 4).
Einen ganz andern Weg, nämlich im Rahmen der sog. *weichen Modellbildung,* haben E. Kofler und G. Menges unlängst beschritten [3]. Ihr Anliegen ist es, flexible, realitätsbezogene Entscheidungsmodelle zu verwenden, die es gestatten, auch *unvollständige* oder *partielle Informationen* optimal einzubeziehen.
Auch dieser Ansatz wird in seinen Grundzügen vorgestellt. (Siehe Fall III im Kapitel 4.)

3. Das Modell der statistischen Entscheidungstheorie

Sei X eine Zufallsvariable, deren Verteilungsfunktion $F(x, z_j)$ vom unbekannten Parameter $z_j \in Z$ abhängt. Die z_j werden als Zustände der Umwelt oder Natur interpretiert und bilden den *Zustandsraum* Z.
Der Statistiker (allgemein der Entscheidungsträger) wählt eine Aktion $a_i \in A$ (oder Letztentscheidung). A heisst *Aktionsraum,* und unter a_i kann man sich z.B. Annahme oder Ablehnung einer statistischen Hypothese vorstellen.

---

2) Sie zerfällt im wesentlichen in Test- und Schätztheorie sowie Versuchsplanung und wurde vor allem von R.A. Fisher, J. Neyman und E.S. Pearson in den zwanziger Jahren unseres Jahrhunderts entwickelt.
3) Der Stichprobenumfang wird nicht zum vornherein fixiert, sondern hängt vom Ergebnis einer Teilstichprobe ab.

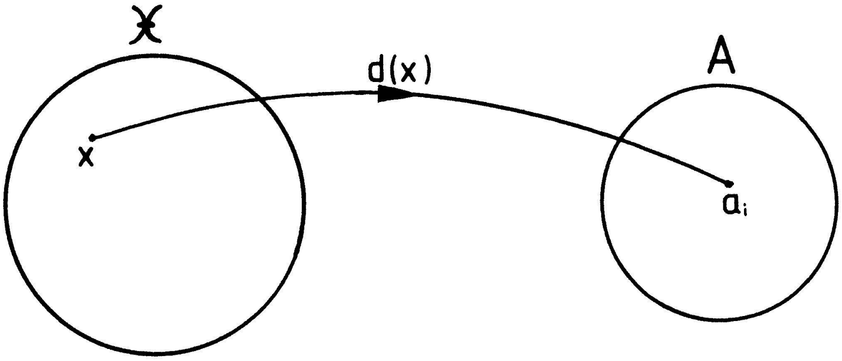
Die ökonomische Bewertung erfolgt über eine *Verlustfunktion*[4]) $v(a_i, z_j) = v_{ij}$, die den Verlust angibt, der bei Verwendung der Aktion a_i im Zustand z_j entsteht. Das Tripel (Z, A, v) beschreibt das ursprüngliche Entscheidungsproblem, das auch matriziell nach (1) dargestellt werden kann.

$$
\begin{array}{c|ccccc}
A \diagdown Z & z_1 & \cdots & z_j & \cdots & z_n \\
\hline
a_1 & v_{11} & & & & v_{1n} \\
\vdots & & & & & \\
a_i & & & v_{ij} & & v_{in} \\
\vdots & & & & & \\
a_k & v_{k1} & & & & v_{kn}
\end{array}
\tag{1}
$$

Die Parameter z_j sind unbekannt und charakterisieren das Verteilungsgesetz $F(x, z_j)$ der Zufallsvariablen X, die wir mittels einer *Stichprobe* beobachten. Damit gewinnen wir zusätzlich Informationen über Z.
Mit $\mathfrak{X}$, dem sog. *Stichprobenraum*, wird die Menge aller Realisationen x bezeichnet, welche X annehmen kann[5]). In Abhängigkeit von $x \in \mathfrak{X}$ hat der Statistiker eine Aktion oder Letztentscheidung $a_i \in A$ zu treffen.

Definition: Unter einer statistischen Entscheidungsfunktion oder Strategie $d(x)$ versteht man eine Abbildung des Stichprobenraumes $\mathfrak{X}$ auf den Aktionsraum A.

(2)



Figur 1

Jeder Strategie $d_i \in D$ (D heisst Strategienmenge) können wir den *mittleren Verlust* oder das *Risiko* $r(d_i, z_j)$ zuordnen, das entsteht, wenn im Zustand z_j die Strategie d_i angewandt wird.
Mit E als Operator der Erwartungswertbildung folgt

$$
\begin{aligned}
r(d_i, z_j) = r_{ij} &= E\{v[d_i(X), z_j]\} \\
&= \sum_{x \in \mathfrak{X}} v[d_i(x), z_j] \cdot p(x \mid z_j).
\end{aligned}
\tag{3}
$$

4) Es ist in der statistischen Entscheidungstheorie üblich, statt eines Gewinnes oder Nutzens $u(a_i, z_j)$ den Verlust $v(a_i, z_j)$ zu verwenden.

5) Zur exemplarischen Verdeutlichung verweisen wir auf das Beispiel in Kapitel 5.

Das Tripel (Z, D, r) stellt das aus (Z, A, v) erweiterte Entscheidungsproblem dar.

$$\begin{array}{cc|ccccc}
 & Z & \varphi_1 & & \varphi_j & & \varphi_n \\
D & & z_1 & \cdots & z_j & \cdots & z_n \\
\hline
\delta_1 & d_1 & r_{11} & & \vdots & & r_{1n} \\
 & \vdots & & & \vdots & & \\
\delta_i & d_i & \cdots & \cdots & r_{ij} & & \\
 & \vdots & & & & & \\
\delta_m & d_m & r_{m1} & & & & r_{mn}
\end{array} \qquad (4)$$

4. Das Fundamentalproblem der statistischen Entscheidungstheorie

Auf der Basis eines *Optimalitätskriteriums* K sowie der Daten $Z, A, v, \mathfrak{X}, r$ und D soll eine im Sinne von *K optimale Strategie* d^* ermittelt werden.
Das Kriterium K hängt vor allem davon ab, wieviel *Information bezüglich der Zustandsmenge* Z vorhanden ist.
In der klassischen Theorie Waldscher Prägung, die ganz im Sinne der harten Modellbildung angelegt ist, werden zwei *Extremfälle* abgehandelt.

Fall I: Es existiert *keine Information* hinsichtlich der Häufigkeit des Auftretens der Zustände z_j. In dieser Situation rät Wald, getragen von einer pessimistischen Grundhaltung, die *Natur* oder Umwelt *als konkurrierenden Gegenspieler* zu interpretieren.
Das Tripel (Z, D, r) wird somit als 2-Personen-Nullsummenspiel des Statistikers gegen die Natur oder Umwelt interpretiert. Damit gelangen die Sätze aus der Spieltheorie sinngemäss zur Anwendung [4], S. 170ff. Wald war sich der Fragwürdigkeit seines pessimistischen Ansatzes bewusst.
Wir gehen jetzt vom Tripel (Z, D, r) nach (4) aus und führen (ganz im Sinne der Spieltheorie) über D und Z Wahrscheinlichkeitsverteilungen $\delta \in \Delta$ und $\varphi \in \Phi$ ein.

$$\varphi = (\varphi_1, \varphi_2, \ldots \varphi_j, \ldots, \varphi_n) \quad \text{mit} \quad \varphi_j \geq 0 \quad \text{und} \quad \sum_{j=1}^{n} \varphi_j = 1$$

heisst *A-priori-Verteilung über Z.*

$$\delta = (\delta_1, \delta_2, \ldots \delta_i, \ldots, \delta_m) \quad \text{mit} \quad \delta_i \geq 0 \quad \text{und} \quad \sum_{i=1}^{m} \delta_i = 1$$

heisst *gemischte Strategie* des Statistikers.
Definition:

$$\bar{r}(\delta, \varphi) = \sum_{i=1}^{m} \sum_{j=1}^{n} r(d_i, z_j) \cdot \delta_i \varphi_j \qquad (5)$$

heisst *mittleres Risiko* oder mittlerer erwarteter Verlust.

Der Statistiker wählt nach dem *Minimax-Kriterium* seine *optimale*, gemischte Strategie δ^*_{I} so, dass das maximale erwartete Risiko minimiert wird, d. h.

$$\min_{\delta\in\Delta} \max_{\varphi\in\Phi} \bar{r}(\delta,\varphi) = \min_{\delta\in\Delta} \max_{z_j\in Z} \bar{r}(\delta,z_j) = \max_{z_j\in Z} \bar{r}(\delta^*_{\mathrm{I}}, z_j). \tag{6a}$$

Analog wird die Natur, welche über φ bzw. z_j verfügt, φ^* (die bezüglich des Statistikers ungünstigste A-priori-Verteilung) so wählen, dass das minimale erwartete Risiko maximiert wird.

$$\max_{\varphi\in\Phi} \min_{\delta\in\Delta} \bar{r}(\delta,\varphi) = \max_{\varphi\in\Phi} \min_{d_i\in D} \bar{r}(d_i,\varphi) = \min_{d_i\in D} \bar{r}(d_i,\varphi^*). \tag{6b}$$

Die optimale Lösung im Fall I führt somit im allgemeinen auf eine gemischte Strategie δ^*_{I} (siehe Figur 3).

Fall II: Wir verfügen über *vollständige Informationen* hinsichtlich der Häufigkeit des Auftretens der z_j, d. h. die Verteilung φ über Z ist bekannt.

$$\bar{r}(d_i,\varphi) = \sum_{j=1}^{n} r(d_i,z_j)\cdot\varphi_j \tag{7}$$

ist das *mittlere Risiko* (auch *Bayes'sches Risiko* genannt), das der Statistiker zu tragen hat, wenn er die Strategie d_i bei Gültigkeit der Zustandsverteilung φ verwendet.
Der Statistiker wählt nach dem Kriterium der *minimalen Verlusterwartung* (Bernoullikriterium) seine optimale Strategie δ^*_{II} bzw. d^*_{II} so, dass

$$\bar{r}(d^*_{\mathrm{II}},\varphi) = \min_{d_i\in D} \bar{r}(d_i,\varphi). \tag{8}$$

Fall III: Partielle Information bezüglich Φ.
Hier betrachten wir die praktisch bedeutsame Situation, in der die *Verteilung* φ über Z *nur teilweise bekannt* ist. Wie in der Einleitung bereits erwähnt, haben Kofler und Menges für die Ausgestaltung dieser Theorie Pionierarbeit geleistet. Im Rahmen dieser Arbeit sollen lediglich die *Grundideen* exemplarisch vorgeführt werden.
Algorithmisch leicht zugänglich und praktisch relevant ist der Fall, wo die partielle Information bezüglich Φ durch ein *System von linearen Ungleichungen* beschrieben werden kann, z. B.
durch Intervallangaben

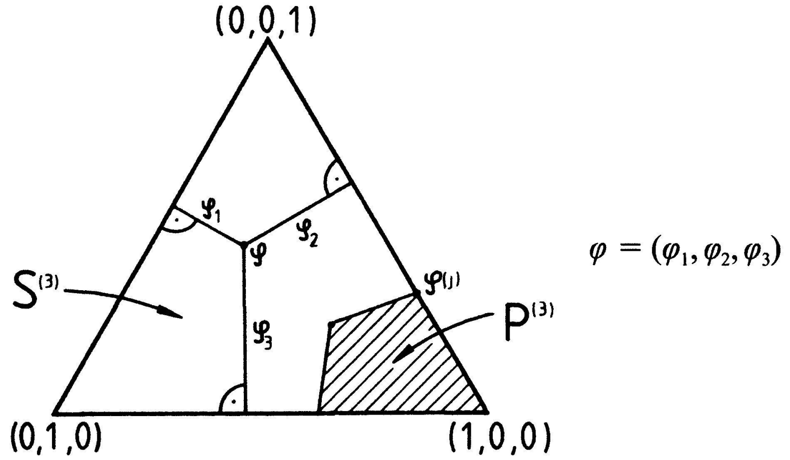
$$a_i \le \varphi_i \le b_i \qquad i = 1,2,3^{6)} \tag{9a}$$

oder durch eine schwache, einfache Ordnung

$$\varphi_1 \le \varphi_2 \le \varphi_3. \tag{9b}$$

6) Wir beschränken uns hier auf den Fall $n = 3$.

Die Menge aller Wahrscheinlichkeitsverteilungen Φ stellt für $n = 3$ das 2dimensionale Simplex $S^{(3)}$ dar, welches in der Figur 2 als gleichseitiges Dreieck mit der Höhe 1 dargestellt ist.



Figur 2

Die linearen Restriktionen in der Form eines Ungleichungssystems führen auf ein konvexes Polyeder $P^{(3)} \subset S^{(3)}$ mit den Eckpunkten bzw. Eckpunkteverteilungen $\varphi^{(j)}$.
Die möglichen A-priori-Verteilungen werden somit auf $P^{(3)}$ eingeschränkt.
Die Entscheidungssituation $(P^{(3)}, D, \bar{r})$ kann wiederum als 2-Personen-Nullsummenspiel des Statistikers (er verfügt über die Strategien $d_i \in D$) gegen die Umwelt (sie wählt eine der unendlich vielen Verteilungen $\varphi \in P^{(3)}$) interpretiert werden.

$$\bar{r}(d_i, \varphi) = \sum_{j=1}^{n} r(d_i, z_j) \cdot \varphi_j. \tag{9c}$$

Der Statistiker wählt nun – ganz im Sinne der Minimax-Philosophie – d^*_{III} bzw. δ^*_{III} so, dass

$$\max_{\varphi \in P^{(3)}} \bar{r}(d^*_{\mathrm{III}}, \varphi) = \min_{d_i \in D} \max_{\varphi \in P^{(3)}} \bar{r}(d_i, \varphi). \tag{10}$$

Es handelt sich bei (10) offenbar um die *Minimierung der maximalen Verlusterwartung* $\bar{r}$, kurz um das Min $\mathrm{E_{Max}}$-Prinzip, der sog. *linearen, partiellen Information,* kurz *LPI* genannt.

Lemma. $\bar{r}(d_i, \varphi) = \sum_{k=1}^{n} r(d_i, z_k) \cdot \varphi_k$ *ist eine lineare Funktion in den* φ_k *und nimmt somit ihr Maximum über dem konvexen Polyeder* $P^{(3)}$ *in einem ihrer Eckpunkte* $\varphi^{(j)}$ *an.*

Das obige Lemma gestattet uns die Reduktion des Spiels $(P^{(3)}, D, \bar{r})$ auf das *endliche* 2-Personen-Nullsummenspiel $(\xi, D, \bar{r})$. Dabei ist ξ die Menge der Eckpunkteverteilungen $\varphi^{(j)}$ und $\bar{r}(d_i, \varphi^{(j)}) = \sum_{k=1}^{n} r(d_i, z_k) \cdot \varphi^j_k$ das erwartete Risiko. $\varphi^j_k (k = 1, 2, \dots n)$ ist die k-te Komponente des Verteilungsvektors $\varphi^{(j)}$.

(10) modifiziert sich in der Form

$$\min_{d_i \in D} \max_{\varphi^{(j)} \in \xi} \bar{r}(d_i, \varphi^{(j)}) = \max_{\varphi^{(j)} \in \xi} \bar{r}(d^*_{\mathrm{III}}, \varphi^{(j)}). \tag{11}$$

Matrizielle Darstellung:
Mit

$$Q_{ne} = (\varphi^j_k) \quad \text{und} \quad R_{\mathrm{mn}} = (r(d_i, z_k)) \tag{11a}$$

folgt

$$R_{mn} \cdot Q_{ne} = L_{me} = (\bar{r}(d_i, \varphi^{(j)})).$$

Das LPI-Spiel wird somit durch die Auszahlungsmatrix L beschrieben:
n Anzahl Zustände z_k der Umwelt,
m Anzahl Strategien d_i des Statistikers,
e Anzahl Eckpunkteverteilungen $\varphi^{(j)}$.

5. Beispiel (Produktionsplanung)

Ein Unternehmen plant die Produktion eines neuen Artikels.

$Z = \{z_1, z_2\}$ $z_1 = 0{,}4$, starke Nachfrage
$z_2 = 0{,}2$, schwache Nachfrage

z_j ist der Anteil der potentiellen Käuferschaft, der das neue Produkt erwirbt.

$A = \{a_1, a_2\}$ a_1 = Produktion einsetzen,
a_2 = auf Produktion verzichten.

Die Verluste $v(a_i, z_j)$ als entgangene Nutzen[7]) interpretiert, setzen wir in der folgenden Matrix fest

$A \setminus Z$	z_1	z_2
a_1	0	100
a_2	40	0

(12)

Es wird nun eine Stichprobe vom Umfang $n = 10$ erhoben[8]), d.h. aus dem potentiellen Käuferkreis werden 10 Personen zufällig ausgewählt und danach befragt, ob sie den geplanten Artikel kaufen werden oder nicht.

7) $v_{ij} = \max\limits_k u_{kj} - u_{ij}$.

8) Diese Zahl wurde so klein gewählt, um einerseits den numerischen Aufwand in vernünftigen Grenzen zu halten und anderseits eine geometrische Interpretation des Problems zu ermöglichen.

$X = X_{10}^{z_j}$, die Anzahl positiver Antworten, ist binominalverteilt, und es gilt

$$p\{X = x|z_j\} = p(x|z_j) = \binom{10}{x} \cdot z_j^x \cdot (1 - z_j)^{10-x}; \qquad x = 0, \ldots, 10. \tag{13}$$

Je nach dem Befragungsresultat x wird die Aktion a_1 oder a_2 ergriffen.
Die Entscheidungsregel oder Strategie d_i wird wie folgt lauten:

$$d_i(x) = \begin{cases} a_1, & \text{wenn } x > i \\ a_2, & \text{wenn } x \le i \end{cases}$$

i heisst Annahmekennzahl.
Gemäss (3) berechnen sich die Risiken $r(d_i, z_j)$

$$r(d_i, z_j) = \sum_{x=0}^{10} v[d_i(x), z_j] \cdot p(x|z_j).$$

Mit Hilfe von (12) und (13) und $F_i^j = \sum_{x=10}^{i} p(x|z_j)$ folgt

$$r(d_i, z_1) = 40 \cdot F_i^1,$$
$$r(d_i, z_2) = 100 \cdot (1 - F_i^2).$$

Die transformierte Entscheidungssituation oder das Spiel (Z, D, r) nimmt dann im Sinne von (4) folgende konkrete Form an[9]):

Z \ D		δ_0	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9
		d_0	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9
φ_1	z_1	0.2	1.9	6.7	15.3	25.3	33.4	37.8	39.5	39.9	40.0
φ_2	z_2	89.3	62.4	32.2	12.1	3.3	0.6	0.1	0.0	0.0	0.0

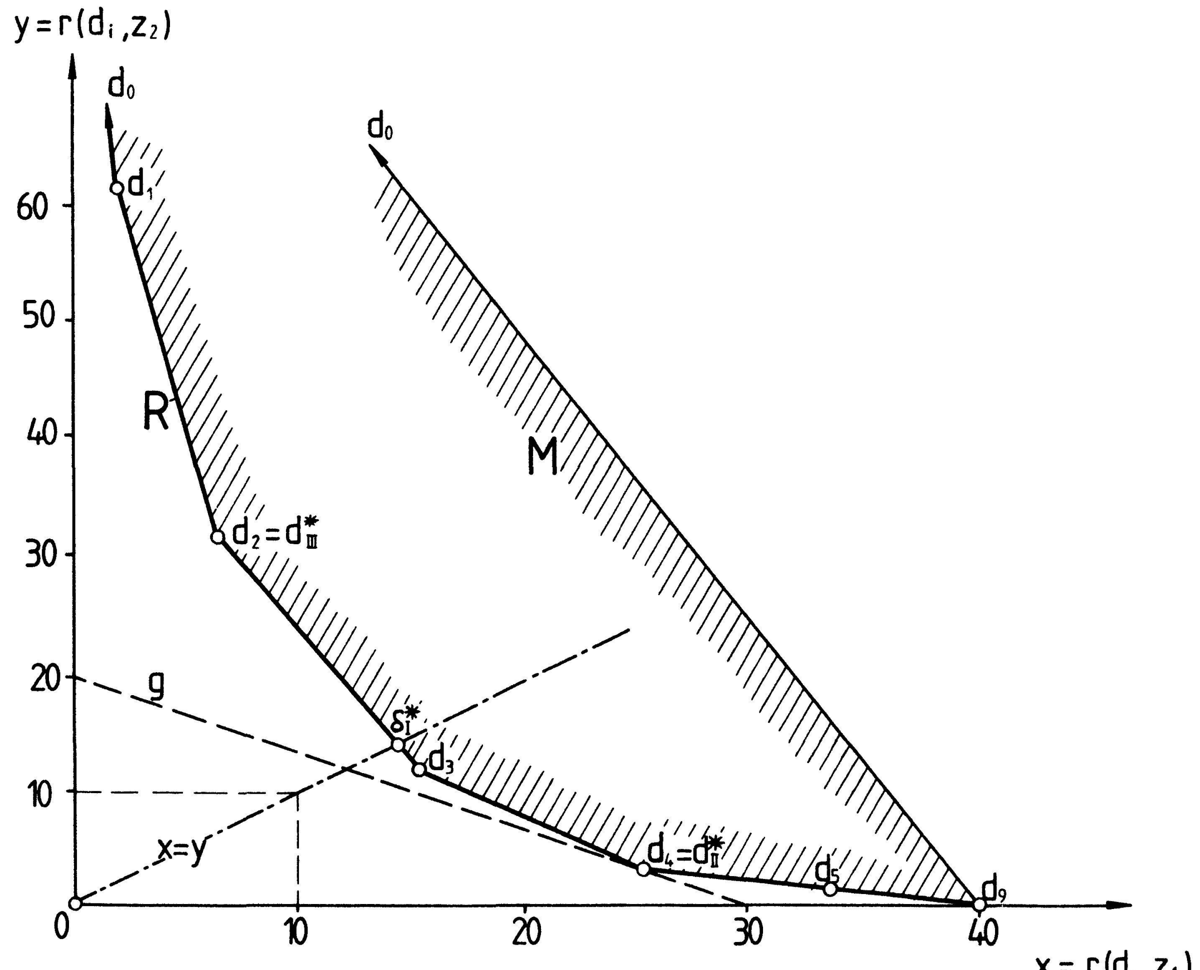
(14)

Da wir nur 2 Zustände z_1 und z_2 haben, lässt sich die obige Risikomatrix R in nachstehender Figur 3 als sog. *Risikomenge M* veranschaulichen, indem man jeder Strategie $d_i \in D$ einen Punkt P mit den Koordinaten $x = r(d_i, z_1)$ und $y = r(d_i, z_2)$ eineindeutig zuordnet.

Fall I (keine Information bezüglich Φ)

Da $\max_{z_j} \min_{d_i} r(d_i, z_j) < \min_{d_i} \max_{z_j} r(d_i, z_j)$, hat das Spiel (14) keine Lösung in *reinen* Strategien $d_i \in D$. Wir gehen deshalb zu gemischten Strategien $\delta \in \Delta$ über mit $\delta = (\delta_0, \ldots, \delta_9)$. $\bar{r}(\delta, z_j)$ $(j = 1, 2)$ sind dann die Koordinaten eines Punktes $P \in M$ der *konvexen Hülle* der 10 Punkte $d_i (i = 0, \ldots, 9)$.

9) Zeilen und Kolonnen sind vertauscht.



Figur 3

Nach (6a) folgt:

$$\min_{\delta} \max_{z_j} \bar{r}(\delta, z_j) = \min_{P \in M} \max(x,y) = \max_{z_j} \bar{r}(\delta_{\mathrm{I}}^*, z_j).$$

Graphisch ergibt sich δ_{I}^* als Schnittpunkt der 1. Winkelhalbierenden $x = y$ mit R, dem sog. *Südwest-Rand* der konvexen Risikomenge M. Die graphische Bestimmung von δ_{I}^* lässt sich nach Figur 3 rechnerisch auswerten und führt auf

$$\delta_{\mathrm{I}}^* = \begin{pmatrix} d_2 & d_3 \\ 0.1115 & 0.8885 \end{pmatrix}.$$

Es sind nur die Strategien d_2 und d_3 zu verwenden, und zwar in der «stochastischen Mischung» 1115:8885.
Der Statistiker wird also im stochastischen Mittel höchstens den Betrag V verlieren, wobei

$$V = \bar{r}(\delta_{\mathrm{I}}^*, z_1) = \bar{r}(\delta_{\mathrm{I}}^*, z_2) = r(d_2, z_1) \cdot 0.1115 + r(d_3, z_1) \cdot 0.8885 = 14.3411.$$

Fall II (vollständige Information bezüglich Φ)
Wir nehmen an $\varphi_1 = \frac{2}{5}$, $\varphi_2 = 1 - \varphi_1 = \frac{3}{5}$, d. h. $\varphi = (\frac{2}{5}, \frac{3}{5})$, es sei also a priori bekannt, dass der

Zustand z_1 (starke Nachfrage) mit einer Wahrscheinlichkeit von $\frac{2}{5}$ eintreffe. Nach (7) berechnet sich das Bayes'sche Risiko

$$\bar{r}(d_i, \varphi) = \frac{2}{5} \cdot r(d_i, z_1) + \frac{3}{5} \cdot r(d_i, z_2) = \frac{2}{5}\, x + \frac{3}{5}\, y .$$

$\bar{r}(d_i, \varphi) =$ konstant bilden eine Schar paralleler Geraden. Unter diesen ist g jene Gerade – sie geht durch d_4 –, welche das kleinste Bayes'sche Risiko trägt. (Siehe Figur 3.)
Somit ist $d^*_{\mathrm{II}} = d_4$.

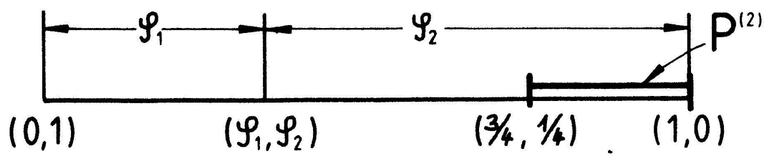
Bemerkung: Die Annahmekennzahl $i = 4$ ist höher als im Fall I, was auch intuitiv leicht einzusehen ist, denn es besteht eine relativ geringe A-priori-Wahrscheinlichkeit $\varphi_1 = \frac{2}{5}$ für starke Nachfrage.
Der erwartete Verlust bei Anwendung von d^*_{II} und bei Gültigkeit von $\varphi = (\frac{2}{5}, \frac{3}{5})$ ist dann

$$\bar{r}(d_4, \varphi) = \frac{2}{5} \cdot r(d_4, z_1) + \frac{3}{5} \cdot r(d_4, z_2) = 12.1 .$$

Fall III (partielle Information bezüglich Φ)
Es sei a priori bekannt, dass starke Nachfrage (z_1) mindestens dreimal so wahrscheinlich sei als schwache Nachfrage (z_2), d.h.

$$\varphi_1 \geq 3 \cdot \varphi_2 ,$$

mit $\varphi_2 = 1 - \varphi_1$ lautet dann die partielle Information: $\varphi_1 \geq \frac{3}{4}$.
Die Menge $\Phi = \{(\varphi_1, \varphi_2) | \varphi_1 + \varphi_2 = 1, \varphi_j \geq 0\}$ wird deshalb eingeschränkt auf $P^{(2)} = \{(\varphi_1, \varphi_2) | \varphi_1 \geq \frac{3}{4}\}$.



Figur 4

Es gibt somit 2 Eckpunktverteilungen $\varphi^{(1)} = (1, 0)$ und $\varphi^{(2)} = (\frac{3}{4}, \frac{1}{4})$. Mit

$$\underset{(14)}{R_{10,2}} = \begin{array}{c} \\ d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_9 \end{array} \begin{array}{c} \begin{array}{cc} z_1 & z_2 \end{array} \\ \begin{pmatrix} 0.2 & 89.3 \\ 1.9 & 62.4 \\ 6.7 & 32.2 \\ \vdots & \vdots \\ 40.0 & 0.0 \end{pmatrix} \end{array} \quad \text{und} \quad Q = \begin{array}{c} \begin{array}{cc} \varphi^{(1)} & \varphi^{(2)} \end{array} \\ \begin{pmatrix} 1 & \frac{3}{4} \\ 0 & \frac{1}{4} \end{pmatrix} \end{array}$$

folgt

$$R_{10,2}\cdot Q_{2,2}=L_{10,2}=\begin{array}{c} \\ d_0\\ d_1\\ d_2\\ d_3\\ d_4\\ d_5\\ d_6\\ d_7\\ d_8\\ d_9\end{array}\begin{pmatrix} \varphi^{(1)} & \varphi^{(2)}\\ 0.2 & 22.5\\ 1.9 & 17.0\\ 6.7 & 13.1\\ 15.3 & 14.5\\ 25.3 & 21.5\\ 33.4 & 25.2\\ 37.8 & 28.4\\ 39.5 & 29.6\\ 39.9 & 29.9\\ 40.0 & 30.0\end{pmatrix}\begin{array}{l} \max\\ 22.5\\ 17.0\\ \boxed{13.1}=\min\max\\ 15.3\\ 25.3\\ 33.4\\ 37.8\\ 39.5\\ 39.9\\ 40.0\end{array} \qquad (15)$$

Nach (11) und (15):

$$\min_{d_i}\ \max_{\varphi^{(j)}} \bar{r}(d_i, \varphi^{(j)}) = \max_{\varphi^{(j)}} \bar{r}(d_2, \varphi^{(j)}) = 13.1,$$

d.h. im Falle der partiellen Information $\varphi_1 \geq \frac{3}{4}$ ist die optimale Strategie $d^*_{\mathrm{III}} = d_2$, und der Statistiker hat höchstens mit einem Risiko von 13.1 zu rechnen.
Kommentar: Die partielle Information $\varphi_1 \geq \frac{3}{4}$ ist optimistisch, und ebenso fällt die entsprechende optimale Strategie aus.

6. Schlussbemerkungen

6.1. Wenn wir die Risikomenge M in Figur 3 betrachten, so fällt auf, dass die Strategien δ auf dem Rand R sog. *zulässig* sind, d.h. es gibt keine andere Strategie von M, die $\delta \in R$ dominiert[10]).
Die Strategien auf dem Rand R bilden eine sog. *minimale vollständige Klasse*[11]) $R \subset M$ von Strategien.
6.2. Die von A. Wald entwickelte statistische Entscheidungstheorie basiert auf dem gleichen Wahrscheinlichkeitsmodell wie die klassische induktive Statistik. Auch Vorentscheidungen mehr oder weniger willkürlicher Art sind bei beiden Modellen notwendig.
Charakteristisch für die entscheidungstheoretische Betrachtungsweise sind
a) der Einbezug einer *ökonomischen Bewertung* richtiger oder falscher Entscheidungen in Form einer Verlustfunktion v;
b) Berücksichtigung von (evtl. subjektiver) *Vorinformation* in Form einer Wahrscheinlichkeitsverteilung φ über dem Zustandsraum Z;
c) Verlagerung des Inferenzproblems auf die Ebene der Optimierung mit *Einschluss spieltheoretischer Aspekte.*

---

10) δ' dominiert δ, wenn $\bar{r}(\delta', z_j) \leq \bar{r}(\delta, z_j)$ für alle j.
11) $R \subset M$ heisst minimal vollständig, wenn zu jeder Strategie $\delta \notin R$ eine Strategie $\delta' \in R$ existiert, die δ dominiert, aber keine Teilmenge von R hat diese Eigenschaft.

Wenn auch der universale entscheidungstheoretische Ansatz sowohl praktisch als auch theoretisch integrierendes Potential besitzt, so sind doch hinsichtlich der Operationalität des Verfahrens einige Vorbehalte zu machen.
Vor allem scheint einerseits die Festsetzung der Verlustfunktion v und anderseits die Frage nach der A-priori-Verteilung φ über dem Zustandsraum etwelche Probleme zu verursachen.
Zu ihrer Lösung scheinen mir deshalb weitere Untersuchungen zum Fall III der partiellen Information (auch hinsichtlich v) erfolgversprechend.

Hans Loeffel
Hochschule St. Gallen

LITERATURVERZEICHNIS

1 J. v. Neumann und O. Morgenstern: Theory of Games and Economic Behavior. Princeton University Press, 1943.
2 A. Wald: Statistical Decision Functions. Wiley, New York 1950.
3 E. Kofler und G. Menges: Entscheidungen bei unvollständiger Information, Nr. 136 der Lecture Notes in Economics and Mathematical Systems. Springer, Berlin 1976.
4 H. Bühlmann, H. Loeffel und E. Nievergelt: Entscheidungs- und Spieltheorie. Springer, Berlin 1975.

 0013-6018/85/060109-12$1.50 + 0.20/0

Kleine Mitteilungen

Ungleichungen für $(e/a)^a\,(b/e)^b$

In einer vor kurzem in dieser Zeitschrift veröffentlichten kleinen Mitteilung [1] sind die beiden Ungleichungen

$$(\sqrt{ab})^{b-a} < \left(\frac{e}{a}\right)^a \left(\frac{b}{e}\right)^b < \left(\frac{a+b}{2}\right)^{b-a}, \qquad b > a > 0, \tag{1}$$

bewiesen worden. (Mit e wird die Eulersche Zahl bezeichnet.) Wird für positive Zahlen a und b für $r \in \mathbf{R}$ der Wert $M_r(a,b)$ durch

$$M_r(a,b) := \begin{cases} \left(\dfrac{a^r + b^r}{2}\right)^{1/r} & \text{für} \quad r \in \mathbf{R} - \{0\} \\ \sqrt{ab} & \text{für} \quad r = 0 \end{cases}$$

definiert, so besagt (1):

$$[M_0(a,b)]^{b-a} < (e/a)^a (b/e)^b < [M_1(a,b)]^{b-a}, \qquad b > a > 0. \tag{2}$$

M_r ist eine bezüglich r in ganz $\mathbf{R}$ stetige Funktion, denn es gilt: $\lim\limits_{r \to 0} M_r(a,b) = \sqrt{ab}$.