3.4 Einige alternative Regeln im Kontext der klassischen Entscheidungstheorie

3.4.1 Die Laplace Regel

Foto:
http://www.mathematik.de/ger/information/landkarte/gebiete/
wahrscheinlichkeitstheorie/wahrscheinlichkeitstheorie.html
[Stand: 25.06.13]

Def. 3.62 (Laplace-Regel)

Gegeben sei das datenfreie Entscheidungsproblem $(\mathbb{A}, \Theta, u(\cdot))$ mit endlichem $\Theta = \{\vartheta_1, \ldots, \vartheta_m\}$.
Die Kriteriumsfunktionen

$$\begin{aligned} \tilde{\Phi} : \quad \mathbb{A} &\longrightarrow \mathbb{R} \\ a &\longmapsto \sum_{j=1}^{m} u(a, \vartheta_j) \end{aligned} \tag{3.15}$$

und

$$\begin{aligned} \Phi : \quad \mathbb{A} &\longrightarrow \mathbb{R} \\ a &\longmapsto \frac{1}{m} \sum_{j=1}^{m} u(a, \vartheta_j) \end{aligned} \tag{3.16}$$

heißen *Laplace-Regel*.

Bem. 3.63

Die beiden Kriteriumsfunktionen (3.15) und (3.16) liefern dieselbe Ordnung auf der Aktionenmenge.

a) Die Kriteriumsfunktion (3.16) entspricht einer Bayes-Regel mit Priori-Verteilung $\pi(\cdot) = (\frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m})$ (mit $m = |\Theta|$), also einer Gleichverteilung auf Θ.
Damit geometrisch: Höhenlinien senkrecht auf Equilisator-Linie.

b) Viele Eigenschaften der optimalen Aktion können deshalb aus den in Kapitel 3.3 formulierten Sätzen über Bayes-Regeln abgeleitet werden. (Zulässigkeit, Entbehrlichkeit randomisierter Aktionen,...)

c) Rechtfertigung durch „Prinzip vom unzureichenden Grund“ (Laplace): Wenn nichts dafür spricht, dass eines der Elementarereignisse wahrscheinlicher ist als die anderen, dann sind sie gleichwahrscheinlich, also

$$\pi(\{\vartheta_1\}) = \pi(\{\vartheta_2\}) = \ldots = \pi(\{\vartheta_m\}).$$

Da

$$\pi(\{\vartheta_1\}) + \pi(\{\vartheta_2\}) + \ldots + \pi(\{\vartheta_m\}) = 1\,,$$

ist zwangsläufig

$$\pi(\{\vartheta_j\}) = \frac{1}{m}\,, \qquad j = 1, \ldots, m.$$

d) Verallgemeinerung auf unendliches Θ:
Theorie der nichtinformativen Priori-Verteilungen, siehe Bemerkung 3.65.

Bem. 3.64 (Beispiel und Kritik)

Abwandlung von Beispiel aus Kapitel 1.3.2, Lotterie
Urnen bestehend aus einer unbekannten Anzahl von grünen, blauen und restlichen (rote, schwarze, violette) Kugeln.
Man kann entweder

a_1 nicht spielen,
a_2 zum Preis von $c_g = 60$€ auf grün setzen oder
a_3 zum Preis von $c_b = 90$€ auf blau setzen.

Es wird eine Kugel zufällig gezogen. Man erhält 240€, wenn die Kugel, auf die man gesetzt hat, gezogen wird.

	$\{g\}$	$\{b\}$	$\{\text{rest}\}$		
a_1	0	0	0	0	
a_2	180	-60	-60	20	$\leftarrow$ optimale Aktion
a_3	-90	150	-90	-10	

Bem. 3.65 („Nichtinformative“ Priori-Verteilung und ihr Informa

In der konditionalen Inferenz (vgl. Kapitel 4.5) gibt es verschiedene Versuche, ähnlich der Laplace-Regel, „nichtinformative“ Priori-Verteilungen zu definieren und diese dann als Standardbewertungen heranzuziehen.

- z.B. die Gleichverteilung, diese ist aber nicht invariant gegenüber Transformationen des Parameters. Man hat dann also „keine Information“ über ϑ, aber eine informative Priori z.B. über eine bijektive, nichtlineare Transformation von ϑ.

- z.B. Verteilungen, die invariant bezüglich bijektiver Transformationen des Parameters sind (Jeffrey-Regel).

- z.B. Verteilungen, die die Entropie maximieren (Jaynes-Regel)

- Ganz neue Möglichkeiten ergeben sich beim Übergang zu Credalmengen (siehe Kapitel 4).

3.4.2 Die Minimax-Regret-Regel von L.J. Savage, auch Niehans-Savage-Regel genannt

Def. 3.66 (Minimax-Regret-Aktion)

Gegeben sei ein datenfreies Entscheidungsproblem $(\mathbb{A}, \Theta, u(\cdot))$ in Nutzenform bzw. $(\mathbb{A}, \Theta, l(\cdot))$ in Verlustform.
Seien – hier der Einfachheit halber – $\mathbb{A}$ und Θ endlich, $\Theta = \{\vartheta_1, \ldots, \vartheta_m\}$, $\mathbb{A} = \{a_1, \ldots, a_n\}$.
Das datenfreie Entscheidungsproblem $(\mathbb{A}, \Theta, r(\cdot))$ in Verlustform mit

$$\begin{aligned} r : \mathbb{A} \times \Theta &\longrightarrow \mathbb{R} \\ (a_i, \vartheta_j) &\longmapsto r(a_i, \vartheta_j) \end{aligned}$$

und

$$r(a_i, \vartheta_j) = \max_{\ell=1,\ldots,n} (u(a_\ell, \vartheta_j)) - u(a_i, \vartheta_j) \tag{3.17}$$

$$\text{bzw.} \qquad r(a_i, \vartheta_j) = l(a_i, \vartheta_j) - \min_{\ell=1,\ldots,n} (l(a_\ell, \vartheta_j)) \tag{3.18}$$

heißt **induziertes Regret Problem** bzw. **induzierte Regret-Tafel**.

Jedes $a^* \in \mathbb{A}$ mit

$$\max_{j=1,\ldots,m} r(a^*, \vartheta_j) \leq \max_{j=1,\ldots,m} r(a, \vartheta_j) \quad \text{für alle } a \in \mathbb{A}$$

heißt **Minimax-Regret-Aktion**.

Bsp. 3.67 (Beispiel und Kritik)

Proposition 3.68 (Bayes-Aktionen in Regrettafeln)

Gegeben sei ein datenfreies Entscheidungsproblem $(\mathbb{A}, \Theta, u(\cdot))$ bzw. $(\mathbb{A}, \Theta, l(\cdot))$ mit $\mathbb{A} < \infty$ und $|\Theta| < \infty$ und die Priori-Bewertung $\pi(\cdot)$. Eine Aktion a^* ist genau dann Bayes-Aktion zu $\pi(\cdot)$, wenn a^* Bayes-Aktion zu $\pi(\cdot)$ im induzierten Regret-Problem ist.

3.4.3 Das Hurwicz-Kriterium

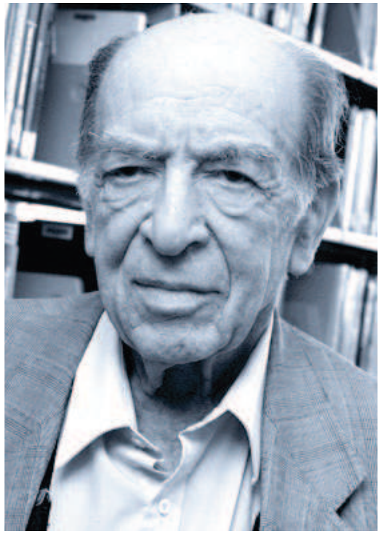


Foto: http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2007/hurwicz-facts.html
[Stand: 25.06.13]

Def. 3.69 (Hurwicz-Kriterium)

Gegeben sei ein datenfreies Entscheidungsproblem $(\mathbb{A}, \Theta, u(\cdot))$ mit $|\Theta| < \infty$, sowie $\alpha \in [0,1]$.

$$\begin{aligned} \Phi : \mathbb{A} &\longrightarrow \mathbb{R} \\ a &\longmapsto \Phi(a) \end{aligned}$$

mit

$$\Phi(a) = \alpha \max_j u(a, \vartheta_j) + (1-\alpha) \min_j u(a, \vartheta_j) \tag{3.19}$$

heißt *Hurwicz-Kriterium* zum *Optimismusparameter* α.

3.4.4 Das Erfahrungskriterium von J.L. Hodges und E.L. Lehmann (1952)

Def. 3.70 (Erfahrungskriterium von Hodges & Lehmann)

Gegeben sei ein datenfreies Entscheidungsproblem $(\mathbb{A}, \Theta, u(\cdot))$, $|\Theta| < \infty$, $\mu \in [0, 1]$ und eine Priori-Bewertung $\pi(\cdot)$.

$$\begin{aligned} \Phi : \mathbb{A} &\longrightarrow \mathbb{R} \\ a &\longmapsto \Phi(a) \end{aligned}$$

mit

$$\Phi(a) = \mu \cdot \left(\sum_{j=1}^{m} u(a, \vartheta_j)\pi(\{\vartheta_j\}) \right) + (1-\mu) \cdot \left(\min_j (u(a, \vartheta_j)) \right) \tag{3.20}$$

heißt *Erfahrungskriterium von Hodges und Lehmann* zum *Vertrauensparameter* μ.

4 Entscheidungsprobleme unter einem allgemeineren Wahrscheinlichkeitsbegriff

4.1 Vorbemerkung: Unsicherheitstheorien

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 became 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".

4.2 Das Ellsberg-„Paradox“ – Vom Ungenügen klassischer Wahrscheinlichkeit

Zur Person: Daniel Ellsberg (geb. 1931 –)

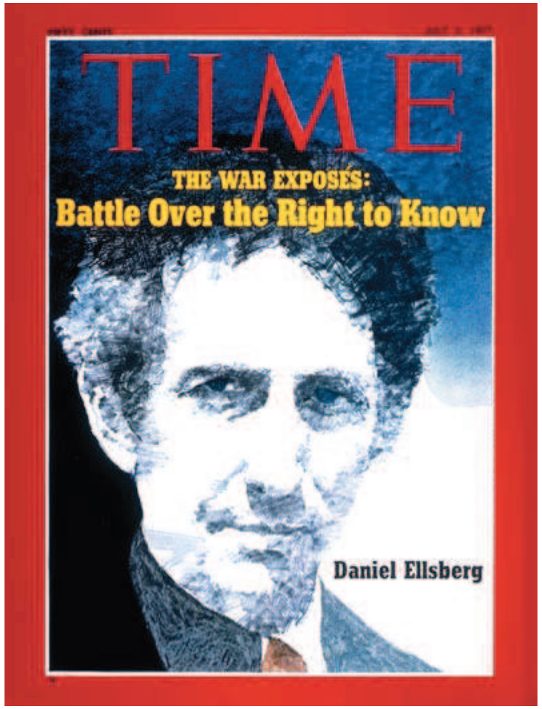


Foto: Time Magazine (July 1971)

Bsp. 4.1 (Das Ellsberg-Experiment)

a) **Das Gedankenexperiment**: Gedankenexperiment in Harvard unter Statistikern und Ökonomen, also Forschern die mit Wahrscheinlichkeitsrechnung vertraut sind, über ihre Präferenzen in folgendem Experiment:

- Gegeben sei eine Urne mit
  r̲oten
  g̲elben Kugeln
  s̲chwarzen

- Anteil der *roten* Kugeln ist genau $\frac{1}{3}$

- Der Anteil der gelben und schwarzen Kugeln ist hingegen unbekannt

- Eine Kugel wird zufällig gezogen

- <u>Situation I</u>:

  * Man darf wählen zwischen

    a_1: 100\$, falls „rot" gezogen wird
    und
    a_2: 100\$, falls „schwarz" gezogen wird

  * was würden Sie bevorzugen?

∗ typische Antwort

$$a_1 \succ a_2$$

- <u>Situation II</u>:

∗ Nun wird gelb zum Joker und man kann wählen zwischen

a_3: 100$; falls „rot" oder „gelb" gezogen wird
und
a_4: 100$; falls „schwarz" oder „gelb" gezogen wird

∗ typische Antwort

$$a_3 \prec a_4$$

b) **Die Priori im Hintergrund**

Laut dem Bayesianischen Paradigma (vgl. Kapitel 3.44) gibt es zu jeder Unsicherheitssituation eine (klassische) Wahrscheinlichkeitsbewertung $\pi(\cdot)$, die die subjektive Einschätzung der Unsicherheit und die daraus abgeleiteten Präferenzen beschreibt. Geht man davon aus, dass jeder Entscheidungsträger seinen subjektiven Erwartungsnutzen maximiert, so kann man $\pi(\cdot)$ durch das beobachtetes Verhalten näher charakterisieren:

Aus Daniel Ellsbergs "Risk, Ambiguity, and the Savage Axioms", p. 655f:

[...] Responses do vary. There are those who do *not* violate the axioms, or say they won´t, even in these situations (e.g., G. Debreu, R. Schlaiffer,

P. Samuelson); such subjects tend to apply the axioms rather than their intuition, and when in doubt, to apply some form of the Principle of Insufficient Reason. Some violate the axioms cheerfully, even with gusto (J. Marschak, N. Dalkey); others sadly but persistently, having looked into their hearts, found conflicts with the axioms and decided, in Samuelson's phrase, to satisfy their preferences and let the axioms satisfy themselves. Still others (H. Raiffa) tend, intuitively, to violate the axioms but feel guilty about it and go back into further analysis.

The important finding is that, after rethinking all their „offending" decisions in the light of the axioms, a number of people who are not only sophisticated but reasonable decide that they wish to persist in their choices. This includes people who previously felt a „first-order commitment" to the axioms, many of them surprised and some dismayed to find that they wished, in these situations, to violate the Sure-thing Principle. Since this group included L. J. Savage, when last tested by me

(I have been reluctant to try him again), it seems to deserve respectful consideration.[...]

Bem. 4.2 (Konsequenzen aus dem Ellsberg-Experiment I)

i) Das Ellsberg-Experiment wendet sich eigentlich ursprünglich gegen das sog. sure-thing-principle, wie es von Savage (1954) zur Fundierung seines Ansatzes verwendet wurde. Dieses Konzept wird in der Vorlesung nicht detailliert betrachtet, entspricht aber letztendlich dem hier verwendeten Bayes-Ansatz.

ii) Bei der in dem Gedankenexperiment von der Mehrheit bevorzugten Entscheidung muss für die zugrunde gelegte Priori-Verteilung also gelten:

$$[\pi(\{r\}) > \pi(\{s\})] \qquad \wedge \qquad [\pi(\{r\}) < \pi(\{s\})]$$

$\Longrightarrow$ Es kann *keine klassische* Wahrscheinlichkeitsverteilung über $(\Theta, \mathcal{P}(\Theta))$ geben, die die am häufigsten beobachteten Präferenzen widerspiegelt.

iii) Man beachte, dass es sich hier um ein *Gedanken*-Experiment handelt. Es geht also nicht darum, dass sich Entscheidungsträger in konkreten Situationen – etwa unter dem Einfluss von Werbung oder Suggestivformulierungen – „irrational"verhalten, sondern darum, dass das oben beschriebene Verhalten von erfahrenen Ökonomen und Statistikern als *rational* empfunden wird. Viele betonen, dass sie bei der Präferenz bleiben, auch wenn sie die Axiome der Wahrscheinlichkeitsrechnung verletzt.

Es gibt also Präferenzen in Entscheidungssituationen unter Unsicherheit, die als *rational* empfunden werden, die aber *nicht* *direkt* mittels *klassischer* Wahrscheinlichkeit beschreibbar sind.

iii) Ein ähnlicher Widerspruch ergibt sich für die am zweithäufigsten gewählte Präferenz $a_1 \prec a_2$ und $a_3 \succ a_4$.

iv) Die Einfachheit des Beispiels stützt das Argument.

Bem. 4.3 (Konsequenzen aus dem E.-Experiment II:)

Nichtadditivität; Ambiguität

- Das zentrale Problem läuft technisch darauf hinaus, dass
$$\pi(\{g,s\}) \neq \pi(\{g\}) + \pi(\{s\}), \tag{4.1}$$
was natürlich das Additionsaxiom für Wahrscheinlichkeiten verletzt.

- „verwerfe solche Wahrscheinlichkeitsbewertungen als irrational“

- versus: „Diese Präferenzen sind aber nicht irrational: Nichtadditivität von Sicherheiten.“

- These: „Es gibt Präferenzen, die – auch nach Reflexion – als rational empfunden werden, aber nicht mittels klassischer Wahrscheinlichkeit beschreibbar sind. Soll die Entscheidungstheorie ihrem Anspruch als Theorie des *rationalen Entscheidens* unter Unsicherheit gerecht werden, so muss sie solche Situationen modellieren können.“

- Das Problem rührt daher, dass sozusagen die Wahrscheinlichkeitsbewertungen „unterschiedlich sicher“ sind.
  Bei $\{g, s\}$ ist die gesamte Unsicherheit auf *Zufälligkeit* reduziert (Anteil $\frac{2}{3}$ bekannt), während bei $\{g\}$ zusätzlich der Anteil unbekannt ist.
  Dieser nichtstochastische Aspekt (*Unbestimmtheit, Ambiguity*) der Unsicherheit stellt ein konstitutives Element der Entscheidungssituation dar und muss geeignet modelliert werden.

- So gilt die Ausgangssituation auch als typisch bei der Konstruktion subjektiver Wahrscheinlichkeiten aus Expertenstatements der Form „ziemlich sicher, dass Krankheit r oder g“, aber „Evidenz nicht aufteilbar zwischen {r} und {g}“.

Bem. 4.4 Exkurs: Neuronale Unterscheidung von Risiko und Ambiguität[6]

	Funktion	Geschwindigkeit
Amygdala	Überwachung	schnell
OFC		
Striatum	Belohnungs- erwartungs- system	nachge- langsam ordnet

[6]Quelle: Hsu, M., et al. (2005, p. 1681)

OFC: Orbitofrontaler Kortex

Experimentelle Behandlungen

- **Kartenstapel**

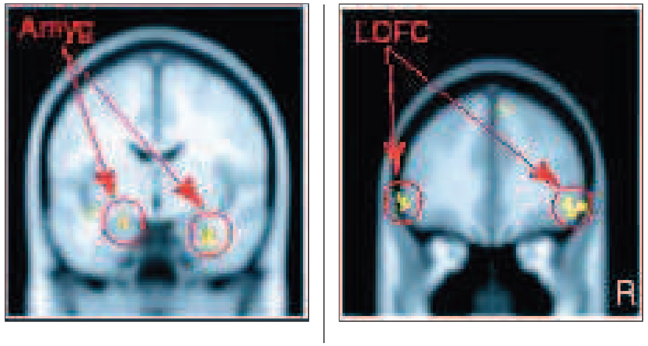
  * Risikobedingung
    Entscheidung zwischen purem Risiko (Zusammensetzung des Kartenstapels bekannt) und sicherem Geldbetrag

  * Ambiguitätsbedingung
    Entscheidung zwischen purer Ambiguität (Zusammensetzung des Kartenstapels unbekannt) und sicherem Geldbetrag

- **Wissen**
  Entscheidungen zwischen sicherem Geldbetrag und Wette auf Bejahung oder Verneinung von Aussagen bzw. Ereignissen, deren Inhalt entweder eine Risiko- oder Ambiguitätsbedingung darstellt.
  Bsp.: Wetter in New York – Wetter in Tirana

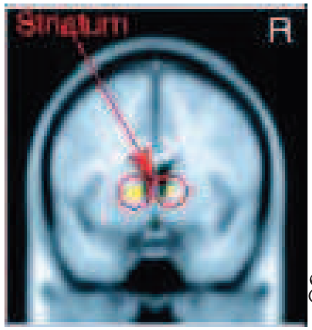
Aktivität von Hirnregionen unter Risiko und Ambiguität

- Während der ambigen Bedingung – im Vergleich zur riskanten Bedingung – waren besonders aktiv *OFC (Orbitofrontaler Kortex) und Amygdala*



[7]

- Während der riskanten Bedingung – im Vergleich zur ambigen Bedingung – war besonders aktiv: *Striatum*



[8]

[7]Quelle: Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., Camerer C.: Neural systems responding to degrees of uncertainty in human decision-making, Science 310, 1681, 2005.

[8]Quelle: Hsu, M., et al. (2005, p. 1682)

<table>
<tr><th></th><th>Funktion</th><th>Geschwindigkeit</th><th>besonders aktiv bei</th></tr>
<tr><td>Amygdala</td><td rowspan="2">Überwachung</td><td rowspan="2">schnell</td><td rowspan="2">Ambiguität</td></tr>
<tr><td>OFC</td></tr>
<tr><td>Striatum</td><td>Belohnungs-
erwartungs-
system</td><td>nachge-
langsam
ordnet</td><td>Risiko</td></tr>
</table>

Bsp. 4.5 (Erste Modellierung des Ellsberg-Experiments)

Erste Schritte zur mathematischen Modellierung:
Die Situation ist beschreibbar:

- einerseits durch *Mengen* klassischer Wahrscheinlichkeitsmaße im Sinne Kolmogorovs. Die Priori-Information besteht aus der Menge aller klassischen[9] Wahrscheinlichkeitsmaße $\pi(\cdot)$ auf $(\{r, g, s\}, \mathcal{P}(\{r, g, s\}))$, mit $\pi(\{r\}) = \frac{1}{3}$ und $\pi(\{g, s\}) = \frac{2}{3}$.

[9]Zur Unterscheidung von intervallwertigen Wahrscheinlichkeiten werden Wahrscheinlichkeitsbewertungen im üblichen Sinn als klassische Wahrscheinlichkeit bezeichnet. Für klassische Wahrscheinlichkeiten werden kleine Buchstaben verwendet, für intervallwertige gro§e Buchstaben.

- andererseits durch *Intervalle*, intervallwertige Wahrscheinlichkeiten:

$$\Pi(\{r\}) = \left[\frac{1}{3}; \frac{1}{3}\right], \ \Pi(\{r, g\}) = \left[\frac{1}{3}; 1\right]$$
$$\Pi(\{g\}) = \left[0; \frac{2}{3}\right], \ \Pi(\{r, s\}) = \left[\frac{1}{3}; 1\right]$$
$$\Pi(\{s\}) = \left[0; \frac{2}{3}\right], \ \Pi(\{g, s\}) = \left[\frac{2}{3}; \frac{2}{3}\right]$$

Die Idee ist verallgemeinerbar!