Ambiguität in (ökonomischer) Entscheidungstheorie

Christian Wurm (Düsseldorf)

July 5, 2023

Hintergrund

Es gibt in den Wirtschaftswissenschaften eine lebhafte Forschung zu Wahrscheinlichkeitstheorie.
Der Grund ist:

- Wahrscheinlichkeitstheorie ist nichts weniger als eine vollständige Theorie der Rationalität
- Ökonomen wollen Wirtschaft verstehen als Interaktion rationaler Agenten
- Insbesondere wollen Sie die *Entscheidungen* rationaler Agenten verstehen bzw. vorhersagen
- Deswegen nennt man das Feld auch **Entscheidungstheorie**

Entscheidungstheorie

Entscheidungstheorie ist ein spannendes Feld, insbesondere weil es zwei Facetten hat:

1. Eine formale Komponente der Theorie der Rationalität, Wahrscheinlichkeit, Spieltheorie etc.
2. Eine empirische Komponente: man überprüft experimentell, welche Entscheidungen Menschen treffen in verschiedenen Szenarien

Dafür werden auch regelmäßig (jährlich) Nobelpreise vergeben, eine nette Lektüre ist Tversky/Kahnemann: Thinking fast and slow.

Ich bin kein Experte in diesem Feld, und spreche nur als Laie!

Entscheidungstheorie

Fishburns Arbeit findet vor diesem Hintergrung statt:

- Es geht darum, Entscheidungen zu erklären und zu verstehen
- und das Mittel der Wahl ist dabei Wahrscheinlichkeitstheorie.

Also schauen wir uns das nochmal an.

Wahrscheinlichkeitstheorie (nach Kolmogorov)

Sei Ω eine erstmal beliebige Menge von **Ergebnissen**. Üblicherweise nimmt man an, es handelt sich um die Ergebnisse eines Zufallsexperimentes.

Ein **Wahrscheinlichkeitsraum** über Ω ist ein Tupel

$$(\Omega, \mathfrak{A}, P)$$

wobei

- $\mathfrak{A} \subseteq \wp(\Omega)$ eine **Boolesche Algebra** ist, und
- $P : \mathfrak{A} \to [0, 1]$ eine Wahrscheinlichkeitsverteilung, die folgende Axiome erfüllt:

1. $P(\emptyset) = 0$
2. $P(\Omega) = 1$
3. falls $A \cap B = \emptyset$, dann $P(A \cup B) = P(A) + P(B)$

Wahrscheinlichkeitstheorie (nach Kolmogorov)

Hieraus folgen alle wichtigen und bekannten Regeln:

$$
\begin{aligned}
P(\overline{A}) &= 1 - P(A) && (1)\\
P(A \cap B) &= P(A)P(A|B) && (2)\\
P(A \cup B) &= P(A) + P(B) - P(A \cap B) && (3)\\
P(B|A) &= P(A|B)\frac{P(B)}{P(A)} && (4)
\end{aligned}
$$

etc.
Wir haben also ein gutes und reiches axiomatisches System. Was bedeutet es aber?

Bedeutung von Wahrscheinlichkeiten

Es gibt zwei verschiedene Interpretationen von Wahrscheinlichkeiten:

Frequentismus: Wahrscheinlichkeit als Grenzwert

In dieser Interpretatation ist die Wahrscheinlichkeit der Grenzwert der relativen Häufigkeit:

$$P(1) = lim_{n\to\infty} \frac{sum_{i=1}^{n} x_i}{n} \tag{5}$$

wobei $x_i \in \{0, 1\}$ das Ergebnis des $i$ten Zufallsexperimentes ist

Das ganze hat verschiedene Probleme:

1. Muss der Grenzwert überhaupt existieren? Rein mathematisch gesehen jedenfalls nicht!

Bedeutung von Wahrscheinlichkeiten

1. Muss der Grenzwert überhaupt existieren? Rein mathematisch gesehen jedenfalls nicht!
2. Für viele Ereignisse ist die Iterierbarkeit konzeptuell nicht mal denkbar; dennoch möchten wir Wahrscheinlichkeiten zuweisen:

(1) Die Wahrscheinlichkeit, dass es im Jahr 2200 noch Menschen auf der Erde gibt, ist gering.

Bedeutung von Wahrscheinlichkeiten

Subjektivismus

In dieser Interpretation ist die Wahrscheinlichkeit die subjektive Sicherheit, die ich einem Eregnis zuweise, z.B.:

- $P(A) = 1$ bedeutet: ich denke A tritt sicher ein
- $P(A) = 0$ bedeutet: ich denke A tritt unmöglich ein
- $P(A) = 0.5$ bedeutet: ich denke nicht dass A sicherer Eintritt als dass es nicht Eintritt
- etc.

Das ist das Credo der Bayesianer. Die Bayesianer haben den Begriff der Wahrscheinlichkeit auch sehr gedehnt, um gewisse Dinge zu beschreiben.

Bayesianische Statistik

Das typische bayesianische Argument hat folgende namengebende Form:

$$P(H|B) = P(B|H)\frac{P(H)}{P(B)} \tag{6}$$

wobei H eine (wissenschaftliche) Hypothese ist, üblicherweise eine Wahrscheinlichkeitsverteilung, und B eine wissenschaftliche Beobachtung.

Ziel wissenschaftlicher Arbeit ist es,

$$argmax_H P(H|B) = P(B|H)\frac{P(H)}{P(B)} = argmax_H P(H|B) = P(B|H)P(H) \tag{7}$$

zu finden. Aber Moment...

Bayesianische Statistik

Moment...

H ist also eine Wahrscheinlichkeitsverteilung P_H, welche wiederum B eine Wahrscheinlichkeit

$$P_H(B)$$

zuweist. Also gilt, nach Kolmogorov, dass P_H, B **im selben Wahrscheinlichkeitsraum** sein müssen. Aber das eine ist ja eine Verteilung über das andere!

Man nennt das auch eine **second order probability**: die Wahrscheinlichkeit einer Wahrscheinlichkeitsverteilung.

Für Bayesianer ist das alles das gleiche, aber die Axiomatik von Kolmogorov gibt das sicher nicht mehr her!

Wahrscheinlichkeiten von Wahrscheinlichkeiten

Das Problem ist allgemeiner: man nehme das folgende

Prinzip der Indifferenz

Gegeben ein Experiment mit $n < \infty$ Ergebnissen, über welches wir keinerlei Wissen haben, nimm an dass alle $\omega \in \Omega$ **dieselbe** Wahrscheinlichkeit haben

$$P(\omega) = \frac{1}{|\Omega|} \tag{8}$$

Das ist nur ein Spezialfall des folgenden Prinzips:

Maximale Entropie

Gegeben ein Experiment über welches wir keinerlei Wissen haben, nimm diejenige Wahrscheinlichkeitsverteilung an, welche die maximale Entropie hat.

Indifferenz

Indifferenz führt aber dazu, das sehr unterschiedliche Szenarien zusammenfallen:

Beispiel 1

Nimm eine Münze oder einen Zufallsgenerator für 0,1. Wir haben $P(1) = 0.5$, und wir sind ziemlich sicher, dass das die reelle (frequentistische) Wahrscheinlichkeit ist.

Beispiel 2

Ich habe die verbliebenen Fusilli und Farfalle zusammengeschüttet. In Ermangelung weiteren Wissens ist unsere Wahrscheinlichkeit, einen Fusillo zu ziehen 0.5. Das ist aber ziemlich sicher nicht die reelle (frequentistische) Wahrscheinlichkeit; wir sind uns diesbezüglich sehr unsicher!

Ambiguität

Wir sehen, dass $P(1) = 0.5$ im Prinzip sehr verschiedene Sachen bedeuten kann – aber das betrifft second order probabilities – und die haben dasselbe Problem!
Das führt uns zum Begriff der **Ambiguität**.

- Der Begriff in dieser Verwendung hängt eng zusammen mit dem Begriff der "Wahrscheinlichkeit zweiter Stufe" –
- soll aber als atomarer, axiomatisch definierter Begriff in Wahrscheinlichkeitsräume eingeführt werden!

Man beachte auch, dass er erstmal nichts mit unsemem linguistischen Begriff zu tun hat, dafür aber – werden wir sehen – mit dem Begriff der Vagheit

Axiomatische Ambiguität

Sei $\mathfrak{P} = (\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum. $a : \mathfrak{A} \to \mathbb{R}_0^+$ ist ein **Ambiguitätsmaß** auf $\mathfrak{P}$, falls es folgende Axiome erfüllt:

A1	$a(\emptyset)$	$= 0 = a(\Omega)$
A2	$a(A)$	$= a(\overline{A})$
A3	$a(A \cup B)$	$\leq a(A) + a(B) - a(A \cap B)$

A4

$a(A) = a(B) \Rightarrow [a(A - B) = a(B - A) \text{ or } a(A \cap B) = a(A \cup B)]$

A5

$a(A) > a(B) \Rightarrow [a(A - B) > a(B - A) \text{ or } a(A \cap B) > a(A \cup B)]$

Axiomatische Ambiguität

Wir sehen also ein Maß, welches in gewissem Sinne orthogonal zu P verläuft:

$$A \subseteq B \not\Rightarrow a(A) \subseteq a(B) \tag{9}$$

In jedem Fall ist es abstrakter:

Fishburn Ambiguität, Intuition

$a(A)$ ist ein Maß dafür, wie **sicher** wir uns sind bezüglich des Urteils $P(A)$. In gewissen Sinne: $a(A)$ ist invers proportional zur "second order" Wahrscheinlichkeit $P_2(\hat{P}(A) = P(A))$, wobei $\hat{P}$ die "wahre" Wahrscheinlichkeit ist. Bayesianisch:

Axiomatische Ambiguität

Fishburn Ambiguität, Intuition

$a(A)$ ist ein Maß dafür, wie **sicher** wir uns sind bezüglich des Urteils $P(A)$. In gewissen Sinne: $a(A)$ ist invers proportional zur "second order" Wahrscheinlichkeit $P_2(\hat{P}(A) = P(A))$, wobei $\hat{P}$ die "wahre" Wahrscheinlichkeit ist.

Je größer die Ambiguität, desto weniger sicher sind wir unserer subjektiven Wahrscheinlichkeit.

$P(A)$ ist die Unsicherheit bezüglich A, $a(A)$ ist die Unsicherheit bezüglich $P(A)$!

Axiomatische Ambiguität

A1 Ist damit klar

A2 Ebenso: die Unsicherheit von $P(A)$ ist gleich der Unsicherheit von $P(\overline{A})$

A3 bedeutet: $a(A)$ wächst **subadditiv** für disjunkte Mengen

Diese Axiome sind **obligatorisch**. A4,A5 sind nicht gleichermaßen unfragwürdig:

A4 bedeutet: falls Mengen gleichermaßen ambig sind, so müssen ihre disjunkten Teile gleichermaßen ambig sein (also ist Ambiguität additiv)

A5 bedeutet dasselbe, nur für die Relation $>$.

Repräsentierbarkeit

Fishburns Frage 1

Gegeben eine Algebra $\mathfrak{A}$ (bei ihm nicht notwendig Boolesch) mit einem Ambiguitätsmaß a, ist es etwas unklar welche Bedeutung eine Zahl $a(A)$ hat:

- Ambiguitäten addieren sich nicht auf 1,
- $a(A) = 0$ hat eine klare Bedeutung,
- $a(A) = 5$ hingegen *nicht*

Es wäre also schön, wenn man von den konkreten Zahlen absehen könnte, und nur eine **partielle Ordnung** von Ambiguitäten aufstellen könnte. Das liefert uns ein System

$$(\mathbf{P}, \succeq)$$

Repräsentierbarkeit

Definition

Ein System $(\mathbf{P}, \succeq)$, $\mathbf{P}$ *ein Wahrscheinlichkeitsraum,* $\succeq$ *eine PO, ist* **repräsentierbar**, *falls gilt: es gibt ein Ambiguitätsmaß* a *auf* $\mathbf{P}$ *so dass gilt:*

$$A \succeq B \iff a(A) \geq a(B) \tag{10}$$

$$A \succ B \iff a(A) > a(B) \tag{11}$$

Die Frage wäre also: welche Eigenschaften muss eine Ordnung haben, um repräsentierbar zu sein?

NB: $\succeq$ kann *nicht* mit $\supseteq$ zusammenfallen!

Repräsentierbarkeit

Das ist relativ geradeaus: $\succeq$ ist eine PO, wobei

$$A \succeq \emptyset \tag{12}$$

$$\Omega \succeq A \tag{13}$$

$$A \succeq \overline{A} \tag{14}$$

$$(A \sim 0, A \cap B = \emptyset) \Rightarrow B \succeq A \cup B \tag{15}$$

Das ist relativ trivial: $\Leftrightarrow$ sorgt ja dafür, dass $\succeq$ (quasi) linear wird.

Repräsentierbarkeit II

Definition

Ein System $(\mathbf{P}, \succeq)$, $\mathbf{P}$ *ein Wahrscheinlichkeitsraum,* $\succeq$ *eine PO, ist* **repräsentierbar II**, *falls gilt: es gibt ein Ambiguitätsmaß* a *auf* $\mathbf{P}$ *so dass gilt:*

$$A \succeq B \Longrightarrow a(A) \geq a(B) \tag{16}$$

$$A \succ B \Longrightarrow a(A) > a(B) \tag{17}$$

Das ist schwächer, aber dadurch eben auch abstrakter und komplexer, da wir mehr Freiheiten haben!

Das wird lange untersucht, aber ganz ehrlich gesagt finde ich das nicht so interessant...

Ambiguität und Produkträume

Zuletzt gibt es noch Ambiguität auf **Produkträumen**. Eine wichtige Bedingung ist dass Ambiguität auf Produkten **additiv** ist.

Das zentrale Ergebnis von Fishburn ist: in diesem Fall kann man die Ambiguität **ausmarginalisieren**. Etwas formaler:

Produktraum

Sei **P** ein Wahrscheinlichkeitsraum über $\Omega_1 \times ...\Omega_n$, und a ein Ambiguitätsmaß auf **P**.
Wir definieren, für $A \subseteq \Omega_i$, $1 \leq i \leq n$,

$$a_i(A) = a(\Omega_1 \times ... \times A \times ... \times \Omega_n) \tag{18}$$

Ambiguität und Produkträume

Theorem

Sei **P** ein Wahrscheinlichkeitsraum über $\Omega_1 \times ...\Omega_n$, und a ein Ambiguitätsmaß auf **P**. Dann gilt: a_i ist ein Ambiguitätsmaß auf Ω_i.

Wenn man weiß, wie wichtig Produkträume in der Wahrscheinlichkeitstheorie sind (Binomialverteilung!), versteht man auch warum dieses Ergebnis zentral ist.

Grundbegriffe

Nehmen wir an, wir haben eine **Familie** von (Wahrscheinlichkeits)verteilungen

$$P_i : \mathfrak{A} \to [0,1] : i \in I$$

die nennen wir $\mathbb{P}$. Wir definieren dann:

$$\lambda_{\mathbb{P}}(A) = min\{P_i(A) : P_i \in \mathbb{P}\} \tag{19}$$

und

$$u_{\mathbb{P}}(A) = max\{P_i(A) : P_i \in \mathbb{P}\} \tag{20}$$

Diese beiden Begriffe haben eine wichtige Bedeutung in der Ökonomie:

Grundbegriffe

Im allgemeinen sind Menschen nicht perfekt rational, insbesondere gibt es eine systematische Divergenz zwischen:

- dem Risiko, Geld/Utilität *nicht zu gewinnen* und
- dem Risiko, Geld/Utilität *zu verlieren*

Kurz gesagt: die meisten empfinden es als viel schlimmer, etwas zu verlieren, als etwas nicht zu gewinnen, selbst wenn es (wie im Fall des Geldes) völlig austauschbar ist!

Diese **Verlustaversion** ist experimentell sehr gut belegt.

Bedeutung von λ, u

Deswegen gibt es folgende intuitive Bedeutung von λ, u.

$\lambda(A)$ Sie sind bereit, jeden Betrag kleiner als $\lambda(A)$ Einheiten zu bezahlen im Austausch dafür, dass Sie eine Einheit *bekommen* falls A eintritt.

$u(A)$ Sie sind bereit, sich für jeden Betrag größer als $u(A)$ Einheiten dafür zu verpflichten, eine Einheit zu bezahlen falls A eintritt.

Man unterscheidet hier also explizit zwischen Gewinn und Verlust!

Man kann λ, u auf diese Art und Weise definieren, ohne jemals auf Wahrscheinlichkeiten zu rekurrieren. Insbesondere können die beide so sein,

- dass wir sichere Verluste einfahren
- dass wir inkonsistent (unlogisch) sind

Gewinn, Verlust, Wahrscheinlichkeit

Folgendes Ergebnis ist zentral für die ökonomische Wahrscheinlichkeitstheorie:

de Finettis Theorem über Wahrscheinlichkeiten

λ, u schützen uns vor sicheren Verlusten (bei iterierem Experiment) und sind konsistent, d.h. $A \subseteq B$ impliziert $\lambda(A) \leq \lambda(B)$, $u(A) \leq u(B)$, gdw. es ein $\mathbb{P}$ gibt so dass $\lambda = \lambda_{\mathbb{P}}$, $u = u_{\mathbb{P}}$.

Das ist deshalb so wichtig, weil es illustriert wie Wahrscheinlichkeiten tatsächlich zentral sind für (ökonomische) Rationalität.

λ, u- Ambiguität

Das Ergebnis sagt auch, dass λ, u nur dann interessant sind, wenn sie wirklich von $\mathbb{P}$ kommen; wir begnügen uns also damit. Die unteren und oberen Schranken führen uns zurück zur Ambiguität:

Theorem 2.3

Definiere das Maß$a : \mathfrak{A} \to [0, 1]$ durch

$$a(A) = u(A) - \lambda(A) \tag{21}$$

Dann erfüllt es A1,A2, aber nicht unbedingt A3-A5.

λ, u

Das liefert uns also nur ein bedingtes Ambiguitätsmaß. Wir erfüllen aber einige praktische Eigenschaften:

$$A \subseteq B \quad \Rightarrow \quad \lambda(A) \leq \lambda(B), \; u \leq u(B) \tag{22}$$

$$A \cap B = \emptyset \quad \Rightarrow \quad \lambda(A \cup B) \geq \lambda(A) + \lambda(B) \tag{23}$$

$$A \cap B = \emptyset \quad \Rightarrow \quad \lambda(A \cup B) \leq \lambda(A) + \lambda(B) \tag{24}$$

$$\lambda(A) + u(\overline{A}) \quad = \quad 1 \tag{25}$$

λ, u- Ambiguität

Wir können 2.3 noch stärken. Sei $\lambda : \wp(\Omega) \to [0,1]$ eine *beliebige* Funktion die erfüllt

1. $\lambda(\emptyset) = 0$
2. $\lambda(\Omega) = 1$
3. $\lambda(A \cup B) \geq \lambda(A) + \lambda(B) - \lambda(A \cap B)$

Wir nennen diese Funktion **konvex**.

Theorem 2.4

Sei λ eine konvexe Funktion, $u(A) := 1 - \lambda(A)$. Dann gibt es $\mathbb{P}$ so dass $\lambda = \lambda_{\mathbb{P}}$, $u = u_{\mathbb{P}}$.

Das nachfolgende ist noch wichtiger:

Theorem 2.5

Sei λ eine konvexe Funktion, $u(A) := 1 - \lambda(\overline{A})$, $a(A) := u(A) - \lambda(A)$. Dann erfüllt a die Axiome A1-A3.

Wir kommen näher!

Belief functions

Der Abschnitt über *belief functions* im Allgemeinen ist relativ kompliziert und nicht besonders interessant für uns.

Interessant ist ein Spezialfall, nämlich **konsonante Glaubensfunktionen**. Wir überspringen die Definition und kommen gleich zum zentralen Ergebnis (welches die komplexe Definition ersetzt):

Theorem 2.5

λ ist eine konsonante Glaubensfunktionen, gdw. gilt $\lambda(\emptyset) = 0$, $\lambda(\Omega) = 1$ und $\lambda(A \cap B) = min(\lambda(A), \lambda(B))$.

Für $u(A) = 1 - \lambda(\overline{A})$ haben wir dann $u(\emptyset) = 0$, $u(\Omega) = 1$ und $u(A \cap B) = max(u(A), u(B))$. Wir nennen u ihr **Konjugat**.

Konsonante Glaubensfunktionen und Vagheit

KGF entste hen noch auf eine andere, natürliche Art und Weise: Sei ϕ ein *fuzy Prädikat* (wie `blau`), mit einer fuzzy Zugehörigkeitsfunktion

$$f_\phi : X \to [0,1]$$

Wir setzen $X := \Omega$, und wir können das Prädikat auf Mengen erweitern mittels

$$\lambda_\phi(A) = min\{f_\phi(x) : x \in A\} \tag{26}$$

$$u_\phi(A) = max\{f_\phi(x) : x \in A\} \tag{27}$$

$$\tag{28}$$

Natürlich gilt dann wieder

$$u_\phi(A) = 1 - \lambda(\overline{A}) \tag{29}$$

Kurz gesagt: wir haben wieder eine KGF

Vagheit und Ambiguität

Das zentrale Ergebnis ist folgendes:

Theorem 4.1

Sei λ eine konsonante Glaubensfunktionen, u ihr Konjungat. Dann gilt: a, definiert durch $a(A) = u(A) - \lambda(A)$ ist ein Ambiguitätsmaß im Sinne von Fishburn (A1-A5).

Und das ist doch ein schönes Ergebnis, weil wir dadurch den etwas abstrakten Begriff mit konkretem Leben füllt. Insbesondere können wir, falls wir a auf diese Art konstruieren, den numerischen Wert

$$a(A) \in [0, 1]$$

sinnvoll interpretieren!