Discrete Choice

Frage: wie wählen Konsumenten aus endlich vielen Alternativen aus?

$a \in A$

Alternative — endliche Menge der Alternativen

① Präferenzen "$\succsim$" auf A, i.e. "$\succsim$" $= R \subseteq A \times A$

$\underbrace{\succsim = R}_{\text{Teilmenge von } A \times A}$

- vollständig : $a \succsim b$ oder $b \succsim a$ $\forall a, b \in A$, $a \neq b$.
- reflexiv : $a \succsim a$ $\forall a \in A$
- transitiv : $a \succsim b$ und $b \succsim c \Rightarrow a \succsim c$ $\forall a, b, c \in A$

Theorem 1: Es existiert $a^* \in A$ mit $a^* \succsim a$ $\forall a \in A$.

Maximales Element

Beweis: Induktion $n = 1$ a_1 - maximal

$A = \{a_1, \dots, a_N\}$

$n \to n+1$ $\underbrace{a_1, \dots, a_n}_{a_i \text{ - sei maximales Element}}, a_{n+1}$

vollständig:

$a_i \succsim a_{n+1}$ — bleibt maximal

$a_i \precsim a_{n+1}$ — wird maximal, da

$\forall j \neq i$, $j = 1, \dots, n$: $a_j \precsim a_i \precsim a_{n+1} \Rightarrow a_j \precsim a_{n+1}$ ($a_j \precsim a_i$: maximal für $a_1, \dots, a_n$; $\Rightarrow$: transitiv) ∎

Theorem: Es existiert $U: A \to \mathbb{R}$ (Nutzenfunktion), so daß $\forall a, b \in A$ $a \succsim b \Leftrightarrow U(a) \geq U(b)$

Beweis: Induktive Konstruktion $A = \{a_1, \dots, a_N\}$

Schritt N: wähle $a_i \in A$ maximal (Theorem 1) und def. $U(a_i) = N$

Entferne a_i aus A und bezeichne $A := A \setminus \{a_i\}$

gehe zu Schritt $N-1$ u.s.w. ∎

Nutzenmaximierung $\max_{a \in A} U(a)$

vollständige Information — perfekte Unterscheidungsfähigkeit

Aber: Auswahlentscheidungen sind oft unsicher und inkonsistent

probabilistischer Ansatz

② Wahrscheinlichkeiten $P_A(a)$, $a \in A$ mit $\sum_{a\in A} P_A(a) = 1$, $P_A(a) \in [0,1]$

W'keit, die Alternative a auszuwählen

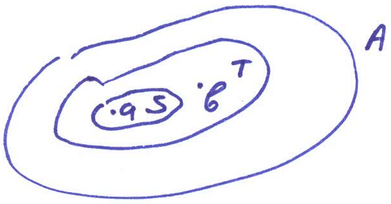
Fehler/Vergessen   ungenügende Information

Bezeichnung: $S \subseteq A$, $P_A(S) = \sum_{a\in S} P_A(a)$ - W'keit, ein Element aus S auszuwählen.

$S, T \subseteq A$, $S \subseteq T$

Axiom (i): $\forall a \in S$: $\underbrace{P_{\{a,b\}}(a) \neq 0,1}_{\text{a und b sind echte Alternativen}}$ $\forall b \in T \Rightarrow \underbrace{P_T(a) = P_T(S) \cdot P_S(a)}_{\text{Pfadunabhängigkeit von S}}$

W'keit a aus T auszuwählen   W'keit S aus T auszuwählen   W'keit a aus S



Axiom (ii): $\underbrace{P_{\{a,b\}}(a) = 0}_{\text{a wird im Vergleich zu b nie gewählt}}$, $a, b \in T \Rightarrow \underbrace{P_T(S) = P_{T\setminus\{a\}}(S\setminus\{a\})}_{\text{a kann entfernt werden}}$

$(S := \{a\} \Rightarrow P_T(a) = P_{T\setminus\{a\}}(\emptyset) = 0)$

Wegen Axiom (ii) betrachte nur $P_{\{a,b\}}(a) \neq 0,1$ $\forall a, b \in A$. (ansonsten W'keit = 0)

Theorem (Luce, 1959): Sei $P_{\{a,b\}}(a|b) \neq 0,1$ $\forall a, b \in A$.

Axiom (i) $\iff \exists u : A \to \mathbb{R}_+$ mit $P_S(a) = \dfrac{u(a)}{\sum_{b\in S} u(b)}$.

und u - eindeutig bis auf positives Multiplikationsfaktor.

Beweis: • Setze $u(a) := \alpha \cdot P_A(a)$, $a \in A$ und $\alpha > 0$ positive Konstante.

Axiom (i) $\underset{T:=A}{\iff}$ $P_A(a) = P_A(S) \cdot P_S(a)$

Also: $P_S(a) = \dfrac{P_A(a)}{P_A(S)} = \dfrac{\alpha P_A(a)}{\sum_{b\in S} \alpha \cdot P_A(b)} = \dfrac{u(a)}{\sum_{b\in S} u(b)}$

• Angenommen u' mit $P_S(a) = \dfrac{u'(a)}{\sum_{b\in S} u'(b)}$

$\Rightarrow u(a) = \alpha \cdot P_A(a) = \dfrac{\alpha u'(a)}{\sum_{b\in A} u'(b)} = \alpha' \cdot u'(a)$ $\forall a \in A$, wobei

$$\alpha' := \frac{\alpha}{\sum_{b\in A} u'(b)} > 0.$$

Interpretation $P_A(a) = \dfrac{u(a)}{\sum_{b\in A} u(b)}$

- Nutzenfunktion $u : A \to \mathbb{R}$: $P_A(a)$ wächst mit Nutzen von a und sinkt mit Nutzen von $b \in A$.

dazu $P_A(a) = \dfrac{u(a)}{u(a) + \sum_{b\neq a} u(b)} = \dfrac{x}{x+y}$, $\left(\dfrac{x}{x+y}\right)'_x = \dfrac{x'\cdot(x+y) - x\cdot(x+y)'}{(x+y)^2} = \dfrac{x+y-x}{(x+y)^2} = \underbrace{\dfrac{y}{(x+y)^2}}_{\text{positiv !}}$

$\left(\dfrac{x}{x+y}\right)'_y = -\dfrac{x\cdot(x+y)'_y}{(x+y)^2} = \underbrace{-\dfrac{x}{(x+y)^2}}_{\text{negativ!}}$

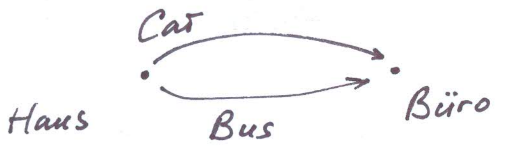
- Independence from Irrelevant Alternatives

$$\frac{P_S(a)}{P_S(c)} = \frac{u(a)}{\sum_{b\in S} u(b)} : \frac{u(c)}{\sum_{b\in S} u(b)} = \frac{u(a)}{u(c)}$$

d.h. $\dfrac{P_S(a)}{P_S(b)} = \dfrac{P_T(a)}{P_T(b)}$ $\forall\, S\subseteq T,\ a, b \in S$

"Verhältnis zwischen den W'keiten für a und b ist unabhängig von der Menge, die a und b enthällt"

Blue/red bus Paradoxon (Debreu, 1960)



$P_{\{C,B\}}(C) = 1/2$, $P_{\{C,B\}}(B) = \frac{1}{2}$, $A = \{C, \overset{\text{blue bus}}{B1}, \underset{\text{red bus}}{B2}\}$

$\underbrace{P_A(B1) = P_A(B2)}$

- gleich wahrscheinlich, ob Car oder Bus
- gleich wahrscheinlich, ob Bus 1 oder Bus 2.

$P_A(C) = \frac{1}{2}$

$\overrightarrow{\text{Intuition}}$ $P_A(B_1) = 1/4 = P_A(B_2)$

Aber Axiom (ii) liefert :

$$\underset{T}{P_A}(\underset{a}{C}) = P_A(\underbrace{C, B_1}_{=S}) \cdot \underbrace{P_{\{C,B_1\}}(C)}_{=\frac{1}{2}}$$

$$\begin{cases} P_A(C, B1) = P_A(C) + P_A(B1) = P_A(C) + \dfrac{1 - P_A(C)}{2} = \dfrac{1}{2}\left(1 + P_A(C)\right) \\ \underbrace{P_A(B_1) + P_A(B2)}_{2 P_A(B_1)} + P_A(C) = 1 \end{cases}$$

$\Rightarrow P_A(C) = \frac{1}{2}\left(1 + P_A(C)\right)\cdot\frac{1}{2} \Rightarrow P_A(C) = \frac{1}{3}$, $P_A(B_1) = P_A(B_2) = \dfrac{1-\frac{1}{3}}{2} = \frac{1}{3}$

Widerspruch zur Intuition.
"ähnliche Alternativen: zu viel W'keit"

· Logit

Transformation / Skalierung

$$w(a) = \ln u(a)$$

Wahrnehmung "psychisch" — Nutzen / Stimulus "physisch"

$$\Rightarrow u(a) = e^{w(a)} \qquad \Rightarrow P_A(a) = \frac{u(a)}{\sum_{b\in A} u(b)} = \frac{e^{w(a)}}{\sum_{b\in A} e^{w(a)}} = \frac{e^{w(a)}}{\sum_{b\in A} e^{w(a)}}$$

Logit model

Weber - Fechner Law aus der Psychophysik:

Frage: wie nimmt man physische Stimuli wahr?

$P(x,y)$ - Wahrscheinlichkeit, daß Stimul x intensiver als y ist.

im physischen Sinne, z.B. lauter, süßer etc.

$\xi_\nu(y) = x \;:\Leftrightarrow\; P(x,y) = \nu$ — W'keitsniveau

Sensitivitätsfunktion: x, die zu y intensiver sind mit W'keit $\nu\in[0,1]$.

$\Delta_\nu(y) = \xi_\nu(y) - \xi_{\frac{1}{2}}(y)$ - Weber-Funktion

Annahme: $\underbrace{\Delta_\nu(y) = y\cdot C(\nu)}_{\text{experimentell getestet}}$ - lineare Weber-Funktion bzgl. y

$\Rightarrow$

- $P(\lambda x, \lambda y) = P(x,y) \;\forall \lambda>0$ - 0-Homogenität.

Sei y fest $\nu_y := P(y,y)$, d.h. $\xi_{\nu_y}(y) = y$, und $\Delta_{\nu_y}(y) = \xi_{\nu_y}(y) - \xi_{\frac{1}{2}}(y)$

Also: $\xi_{\frac{1}{2}}(y) = y - y\cdot C(\nu_y) = y(1 - C(\nu_y))$.

$P(\lambda x, \lambda y) = \nu \Leftrightarrow \xi_\nu(\lambda y) = \lambda x \Leftrightarrow \frac{1}{\lambda}\xi_\nu(\lambda y) = x$

Aber: $x = \frac{1}{\lambda}\xi_\nu(\lambda y) \underset{\text{def.}}{=} \frac{1}{\lambda}\left(\Delta_\nu(\lambda y) + \xi_{\frac{1}{2}}(\lambda y)\right) = \frac{1}{\lambda}\left(\lambda y\cdot C(\nu) + \lambda y(1 - C(\nu_{\lambda y}))\right)$

$\nu_{\lambda y} = P(\lambda y, \lambda y) = P(y,y)$ "Nur Differenz wichtig!"

$\underbrace{y\,C(\nu)}_{\Delta_\nu(y)} + \underbrace{y(1 - C(\nu_y))}_{\xi_{\frac{1}{2}}(y)} \underset{\text{def.}}{=} \xi_\nu(y) \Leftrightarrow P(x,y) = \nu.$

- $P(x,y) \underset{\lambda := \frac{1}{y}}{=} P\left(\frac{x}{y}, \frac{y}{y}\right) = P\left(\frac{x}{y}, 1\right) = P\left(e^{\ln x - \ln y}, 1\right) = H(\ln x - \ln y)$

$H(s) = P(e^s, 1)$ - monoton steigend

$w(x) := \ln(x)$ Wahrnehmungsfunktion

Fechner-Law

$P(x,y) = H(\ln x - \ln y)$