

UTILITY THEORY

リスク認知と効用理論

伊庭 斉志

St. Petersburgのパラドクス

理想的なコインを表がでるまで投げ続ける。1回目で出れば2円、2回目で出れば4円、3回目で出れば8円、...、 n 回目なら 2^n 円を受け取る。

この賭けで得られる金額の期待値はいくらだろうか？

いくらに参加料ならこの賭けに参加すべきか？

Toss	Winning
H	\$2
TH	\$4
TTH	\$8
TTTH	\$16
...	...

St. Petersburg のパラドクス

期待値の計算

$$\langle \text{Winning} \rangle_N = \sum_{k=1}^N P(k)W(k) = \sum_{k=1}^N \frac{2^k}{2^k} = \sum_{k=1}^N 1 = N$$

$$\lim_{N \rightarrow \infty} \langle \text{Winning} \rangle_N = \infty$$

この賭けの価値は本当に無限大か？

何かが変だ？？？

Bernoulliの登場



Daniel Bernoulli(1700-1782): お金の効用は、お金の金額そのものではない。

お金の効用 $=\log(\text{金額})$ のようなもの

$$\langle \text{Utility} \rangle_N = \sum_{k=1}^N P(k) U(k) = \sum_{k=1}^N \frac{\ln(2^k)}{2^k} = \ln(2) \sum_{k=1}^N \frac{k}{2^k}$$

$$= 2\ln(2) \left[1 - \left(\frac{1}{2}\right)^N - N \left(\frac{1}{2}\right)^{N+1} \right]$$

$$\lim_{N \rightarrow \infty} \langle \text{Utility} \rangle_N = 2\ln(2) = \ln(4)$$

金額=4円の
価値しかない

効用とは何か？

- 効用 (utility) : 個人が多数の行動の1つを選択するとき、それを最大にしたいと望むもの
 - フォン・ノイマン、モルゲンシュタインらによる効用理論
- 相対的な関係からどちらかを選択するか、として効用を定義する
 - 順序のみが重要、基準点はなく、数量的でもない
- 選好 (preference) : 選択の結果
- 効用関数 : 選好関係を数量的な関数で表したもの

Von Neumann, J. and Morgenstern, O. (1947)
Theory of Games and Economic Behavior. Princeton
University Press: Princeton.

効用関数 $U(S)$

- 異なる状態 S 間の選好は効用関数 $U(S)$ で捉えられる。
- $U(S_i) > U(S_j)$ なら、 S_i を S_j より好む
-
- $U(S_i) = U(S_j)$ なら、2つの状態 S_i と S_j の違いは感じない(無差別である)

期待効用の最大化

合理的な人間は、期待効用 (EU) を最大にする行動を選択する。

$$E(X) = \sum_{i=1}^n \pi_i U(x_i)$$

$$EU(A | \mathbf{E}) = \sum_i P(\text{Result}_i(A) | \text{Do}(A), \mathbf{E}) U(\text{Result}_i(A))$$

ここで $\text{Result}_i(A)$ は、行動 A を実行したあとの可能なすべての状態を表す

効用理論の基礎

記法:

$A \succ B$ A が B より好まれる

$A \sim B$ A と B は無差別である

$A \succeq B$ A が B より好まれるか、もしくは2つは無差別である

6つの公理

Orderability	$(A \succ B) \vee (B \succ A) \vee (A \sim B)$	You must make a decision
Transitivity	$(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$	
Continuity	$(A \succ B \succ C) \Rightarrow \exists p [p, A; 1-p, C] \sim B$	
Substitutability	$(A \sim B) \Rightarrow [p, A; 1-p, C] \sim [p, B; 1-p, C]$	
Monotonicity	$(A \succ B) \Rightarrow (p \geq q \Leftrightarrow [p, A; 1-p, B] \succeq [q, A; 1-q, B])$	
Decomposability	$[p, A; 1-p, [q, B; 1-q, C]] \sim [p, A; (1-p)q, B; (1-p)(1-q), C]$	

この公理が満たされてるときには、以下の条件をみたすような実数値効用関数 U が存在する。

$$U(A) > U(B) \Leftrightarrow A \succ B$$

$$U(A) = U(B) \Leftrightarrow A \sim B$$

von Neumann-Morgenstern

合理的なエージェントの公理

- 推移性: A と B、および B と Cが無差別であるなら、A と C も無差別でなくてはならない。
- 連続性: A がB より好まれるなら、ある確率 α があつて、 αA と Bは無差別である。

von Neumann-Morgenstern 合理的なエージェントの公理

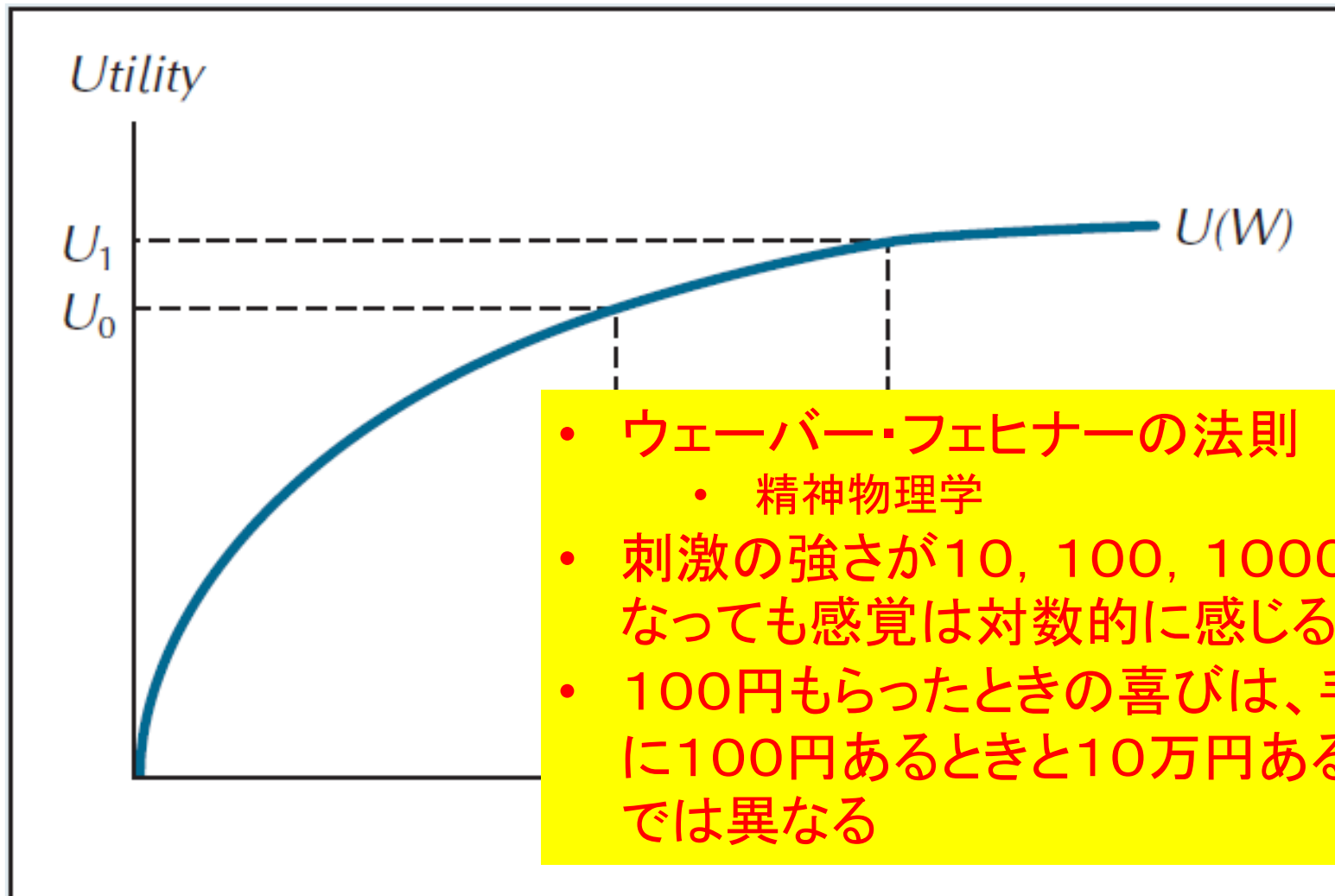
- 独立性: A と B が無差別なら、どのような確率でも、 αA と αB は無差別である。
- *Desire for high probability of success*: $\alpha_1 > \alpha_2$ なら $\alpha_1 A$ は $\alpha_2 A$ よりも好まれる。
- 合成確率: αA と B が無差別で、 $\alpha = \alpha_1 \alpha_2$ なら、 $\alpha_1 \alpha_2 A$ と B は無差別である。

無関係な選択肢からの独立性

- **推移性**: A と B、および B と C が無差別であるなら、A と C も無差別でなくてはならない。
- 2つの選好順位はそれのみで決まり、他の要素に影響されない。
- 本当か？

効用関数の例

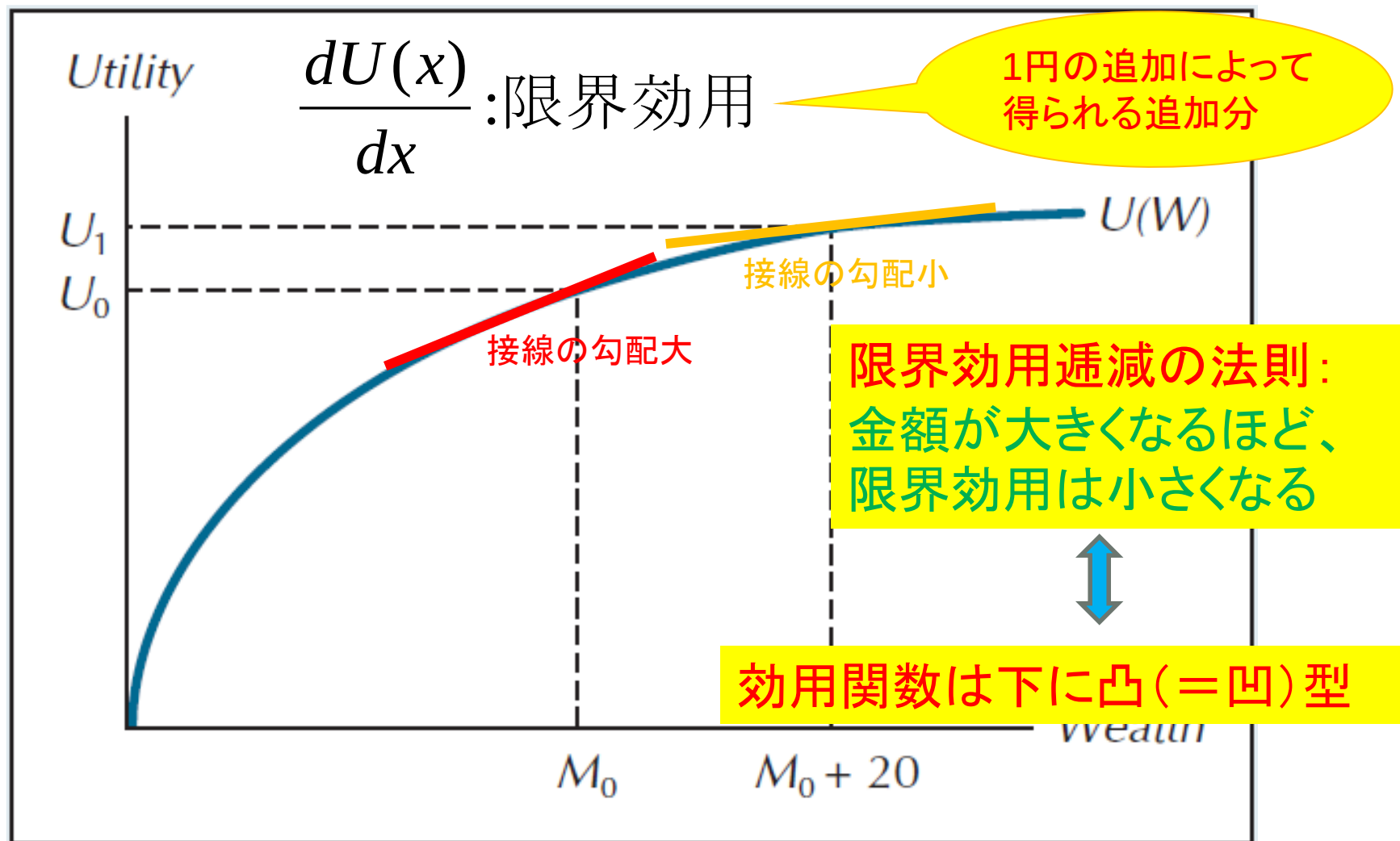
$$U(x) = k \log(x) \quad k > 0$$



- ウェーバー・フェヒナーの法則
 - 精神物理学
- 刺激の強さが10, 100, 1000倍となっても感覚は対数的に感じる
- 100円もらったときの喜びは、手元に100円あるときと10万円あるときでは異なる

効用関数の例

$$U(x) = k \log(x) \quad k > 0$$



効用関数の他の例

$$U(x) = 1 - e^{-x/\lambda} \quad \lambda > 0$$

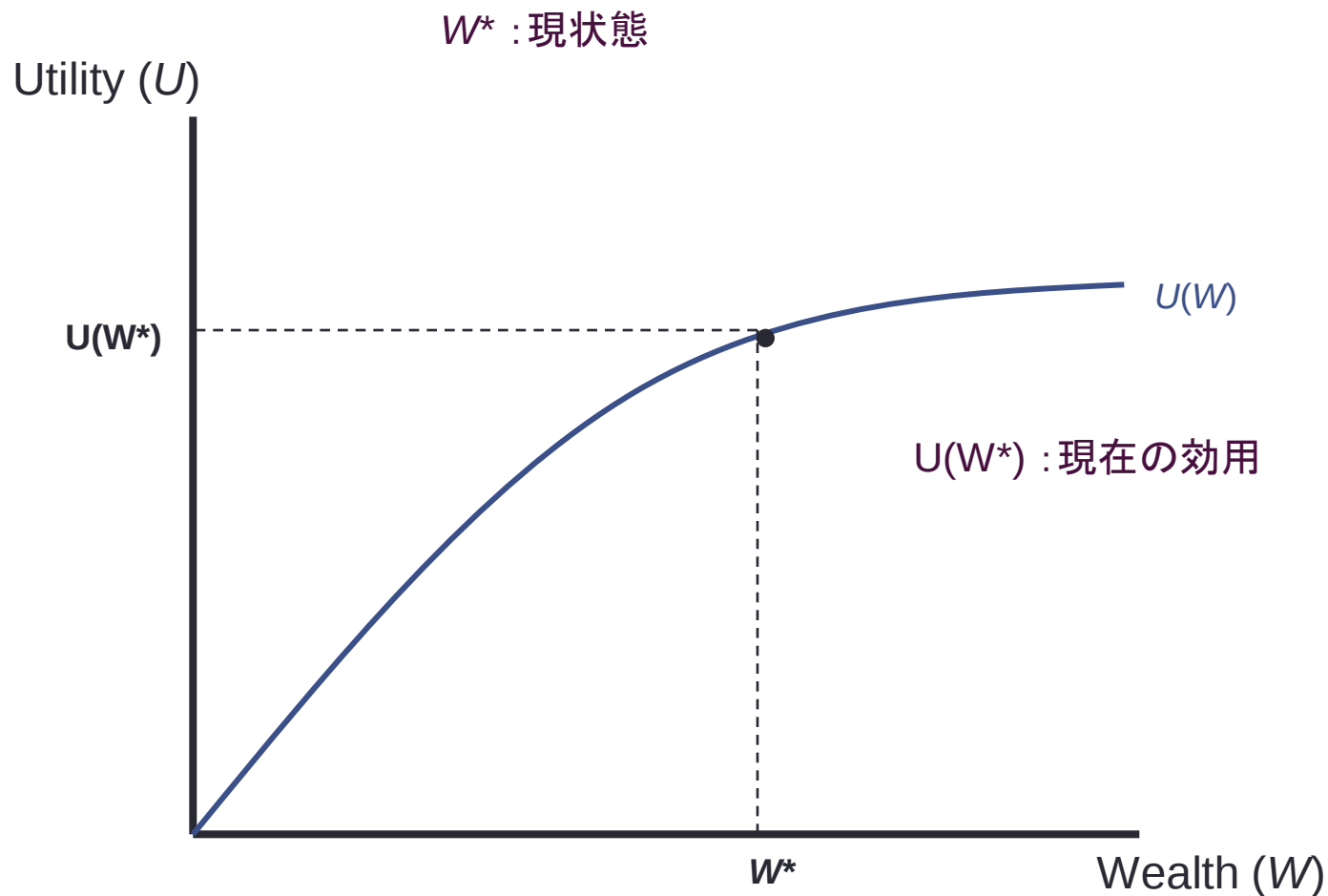
下に凸(=凹)型
限界効用逓減

$$U(x) = \begin{cases} \sqrt{x} & x \geq 0 \text{のとき} \\ -x^2 & x < 0 \text{のとき} \end{cases}$$

限界効用逓減ではない

リスク回避

効用関数が下に凸(=凹)型するとき



リスク回避

効用関数が下に凸(=凹)型するとき

- 次のような賭けを考えよう:

- (a) 五分五分の確率で、\$ h を得るか、失う

$$U^h(W^*) = \frac{1}{2} U(W^* + h) + \frac{1}{2} U(W^* - h)$$

- (b) 五分五分の確率で、\$ $2h$ を得るか、失う

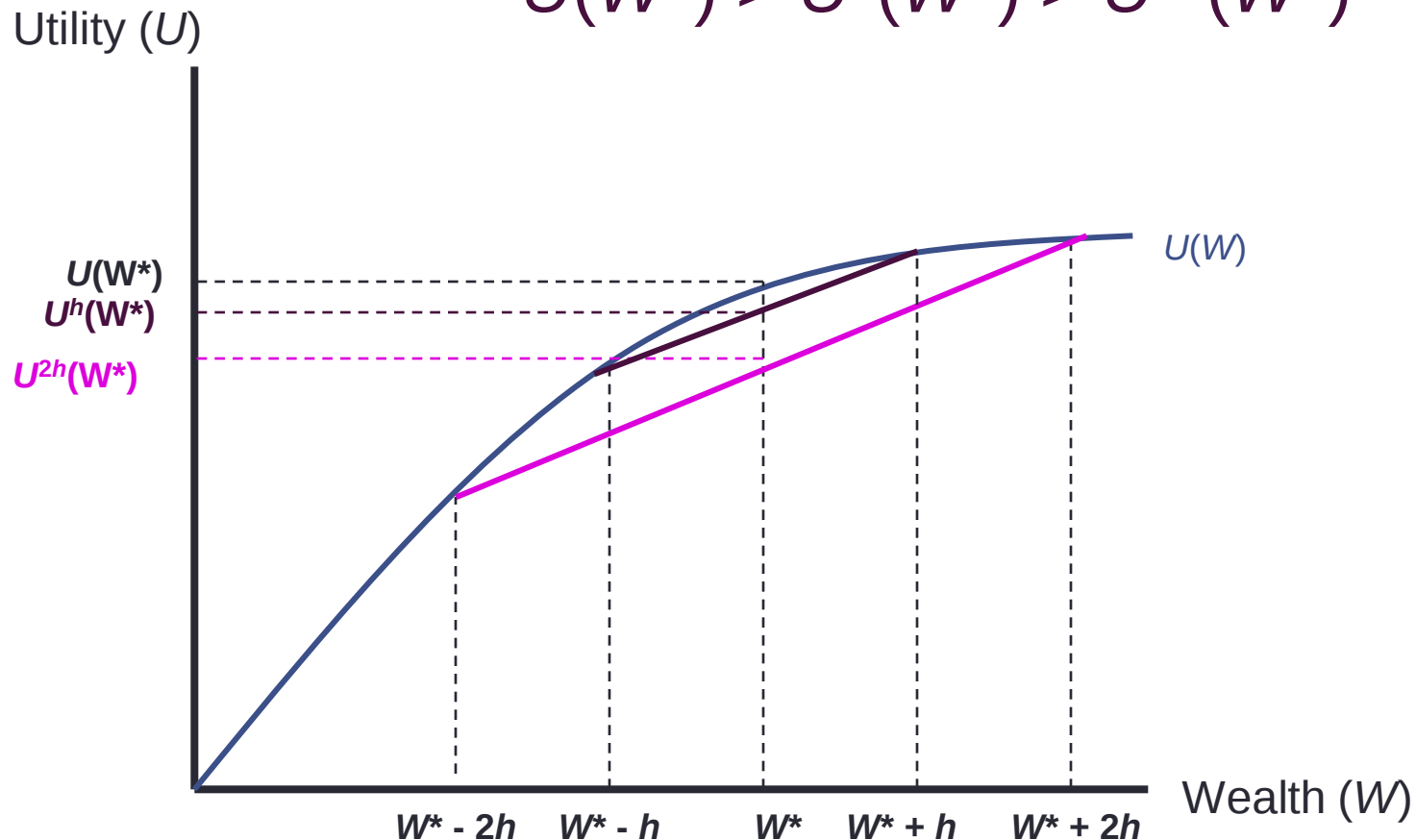
$$U^{2h}(W^*) = \frac{1}{2} U(W^* + 2h) + \frac{1}{2} U(W^* - 2h)$$



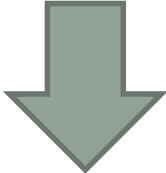
リスク回避

効用関数が下に凸(=凹)型するとき

$$U(W^*) > U^h(W^*) > U^{2h}(W^*)$$

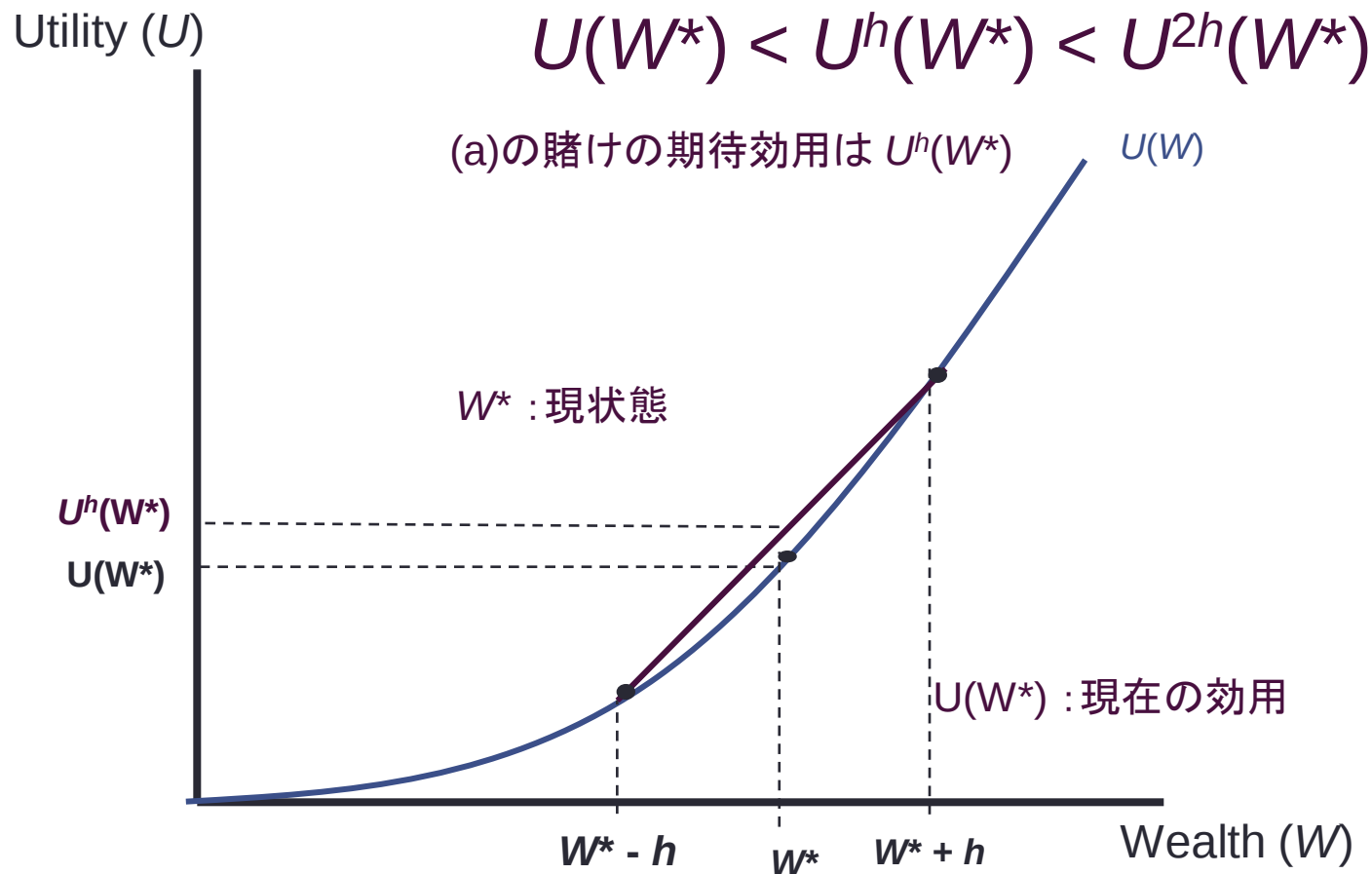


リスク回避

- 効用関数が下に凸(=凹)型するとき
 - 賭けがあるような状態よりも、現状を好む
 - 大きな賭けよりも小さな賭けのほうを好む
- 
- 賭けを避けるために少額をはらうことがある
 - なぜ人は保険にはいるのか？

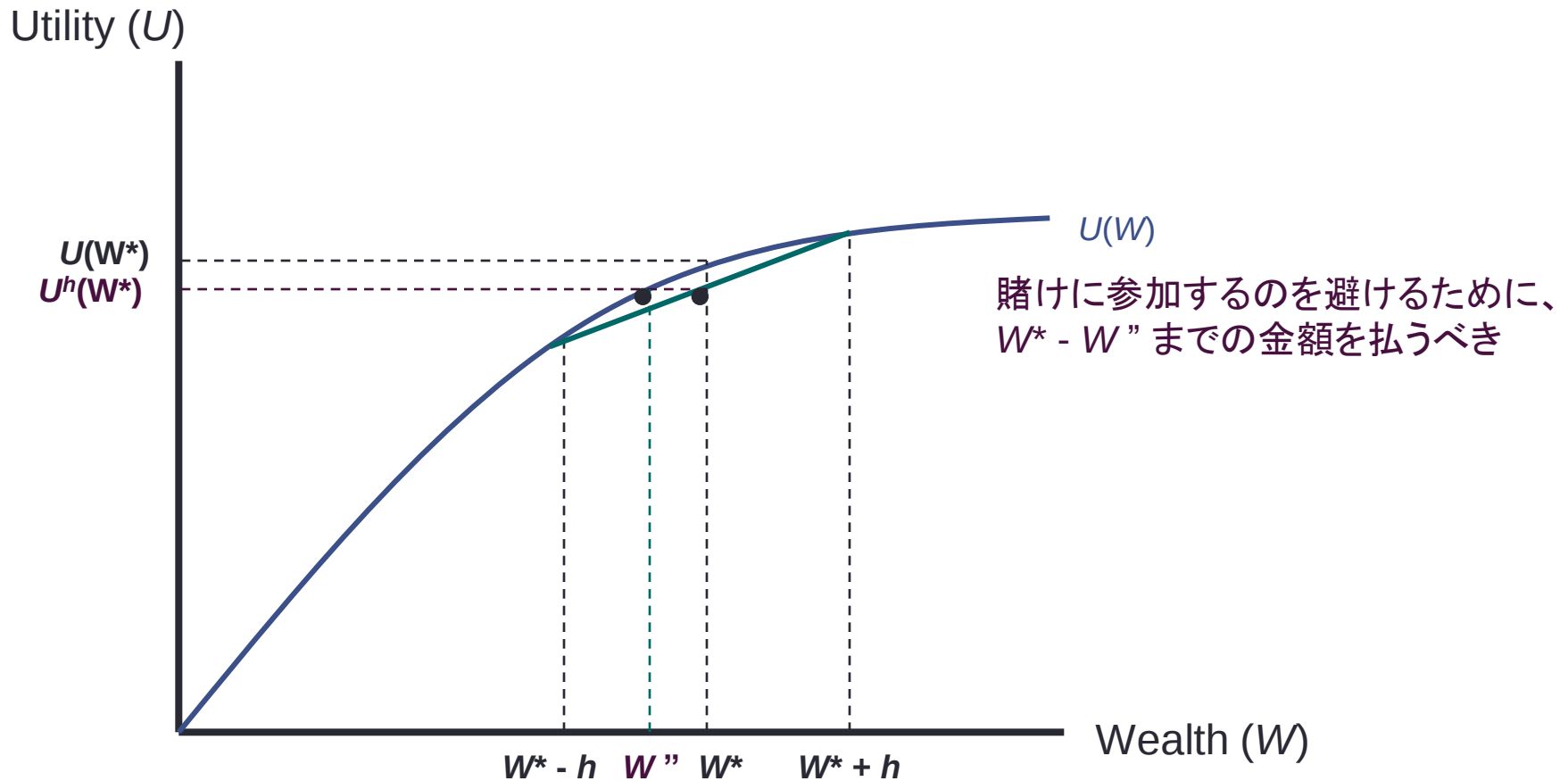
リスク追及

効用関数が上に凸型するとき

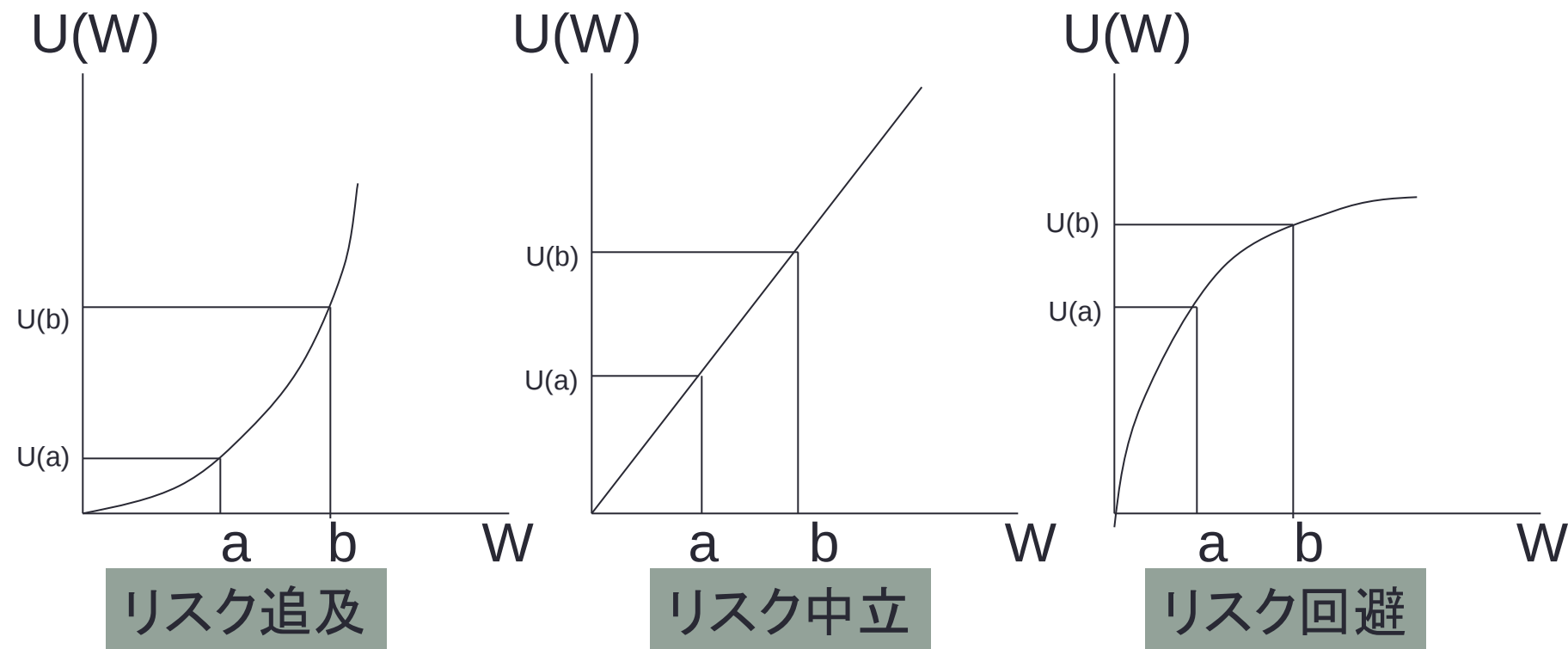


リスク回避と保険

W'' : 賭け(a)と同じ状態



リスクを好むか、嫌うか？



Goals:

- (a) Formally define what is risk-aversion.
- (b) Establish the concept of risk-premium

リスク回避と保険

- リスク回避＝(公平な)賭けを常に避けること
- この場合、限界効用逓減となる
- また、賭けを避けるために金額をはらう

なぜ保険にはいるのか？

- \$100,000 の財産を持っている人を考える
- この人が、\$20,000 の車を購入した
- 25% の確率で事故・盗難にあうとする
- この人は、von Neumann-Morgenstern 効用指数に従うとする

$$U(W) = \ln (W)$$

なぜ保険にはいるのか？

- 期待効用を求めると、

$$E(U) = 0.75U(100,000) + 0.25U(80,000)$$

$$E(U) = 0.75 \ln(100,000) + 0.25 \ln(80,000)$$

$$E(U) = 11.45714$$

- この状況では、公平な保険のプレミアムは、\$5,000 (\$20,000の25%)である

なぜ保険にはいるのか？

- 賭けをさけるためには、\$5,000以上でもはらうだろう。
- いくらまでならはらうか？

$$E(U) = U(100,000 - x) = \ln(100,000 - x) = 11.45714$$

$$100,000 - x = e^{11.45714}$$

$$x = 5,426$$

- 最大のプレミアムは、\$5,426 である

古典的な効用理論への批判

- リスクは単一の数（期待効用）ではとらえきれない
- 効用が多くの人にとって意味をなさない
- 自然な効用関数は存在しない
 - どんなときに対数や平方根をもちいるべきか？
- 人は公理に従わない
 - Prospect Theory from Daniel Kahneman and Amos Tversky

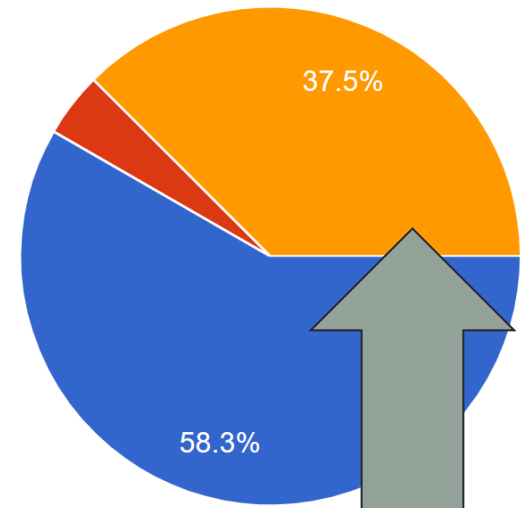
新聞購読の値段

おとり広告

質問3：あなたは自分が大好きな雑誌や新聞の年間購読をしようとしていて、
の選択肢があるとしよう。このうちどれを好むか？ できるだけ直感的に

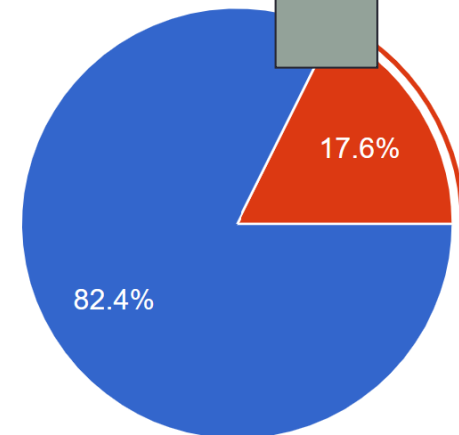
- ☐ 一年間自由にアクセスできるウェブ版WWWのみの購読 600円 (毎月あたり)
- ☒ 一年間分の印刷版のみの購読 1,200円 (毎月あたり)
- ☐ 一年間分の印刷版とウェブ版WWWのセット購読 1,200円 (毎月あたり)

17.6%が37.5%に増えている!!
これは不合理だ



質問3：あなたは自分が大好きな雑誌や新聞の年間購読をしようとしていて、
の選択肢があるとしよう。このうちどれを好むか？ できるだけ直感的に

- ☐ 一年間自由にアクセスできるウェブ版WWWのみの購読 600円 (毎月あたり)
- ☒ 一年間分の印刷版とウェブ版WWWのセット購読 1,200円 (毎月あたり)

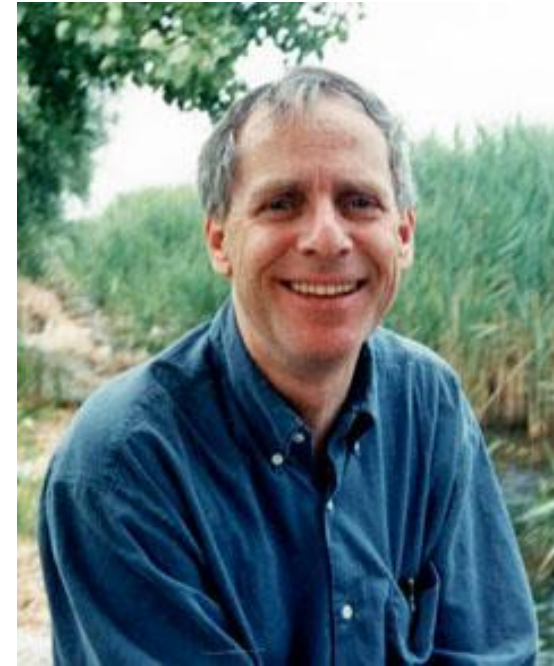
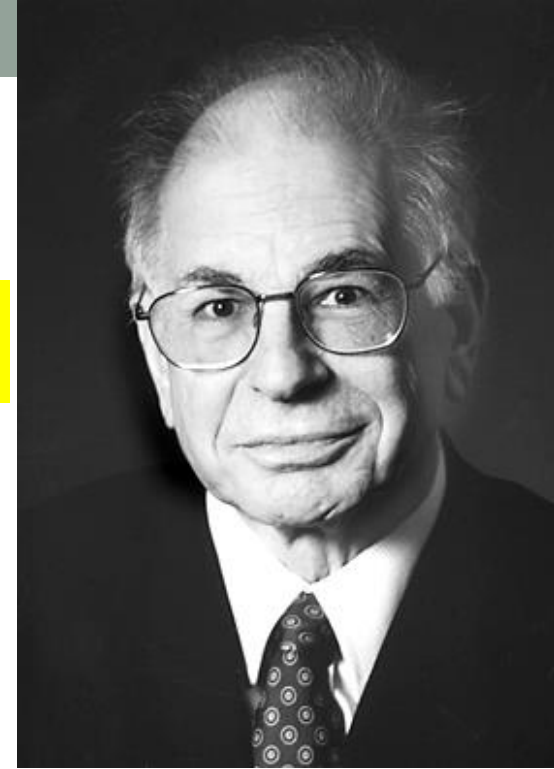


プロスペクト理論

ダニエル・カーネマン(1934年-)
2002年: ノーベル経済学賞

- 行動経済学
- 心理学＋経済学
 - *Prospect Theory: An Analysis of Decision under Risk*, by Daniel Kahneman and Amos Tversky, *Econometrica*, Vol. 47, No. 2 (March 1979), pp. 263-292
- 合理的なエージェント(エコノ)は存在しない
- Key concepts
 - 損失回避
 - 現状バイアス
 - 授かり効果(保有効果)

エイモス・トベルスキー
(1937年 - 1996年)



古典的効用理論批判:アレのパラドクス



Xの期待効用を U_X とする。
 $U(500万)=1$, $U(100万)=\alpha$, $U(0)=0$ としよう。
すると、
 $U_A=\alpha$, $U_B=0.1 \times 1 + 0.89\alpha$, $U_C=0.11\alpha$, $U_D=0.1$
よって、

$$A \geq B \Leftrightarrow U_A \geq U_B \Leftrightarrow 0.11\alpha \geq 0.1$$

$$C \geq D \Leftrightarrow U_C \geq U_D \Leftrightarrow 0.11\alpha \geq 0.1$$

なので、 $A \geq B$ かつ $D \geq C$ なら矛盾する。

期待的効用理論によると、以下のようにしなければならない

- Aを選んだ人はCを選ぶ
- Bを選んだ人はDを選ぶ

古典的効用理論批判: アレのパラドクス

- フランスの経済学者 M.Allaisによる反例

期待的効用理論によると、以下のようにならない

- Aを選んだ人はCを選ぶ
- Bを選んだ人はDを選ぶ

- しかしながら、アレは多くのヒトは、 $A > B$ かつ $D > C$ であると予想した



エイモス・トベルスキー
(1937年 - 1996年)

- Tverskyによる実験結果
 - 29名のうち17名がアリの予想通りの選好パターンを示した
 - そのあとで、期待効用理論について解説したあと再び意見をもめたが、17名全員が意見を変えなかった

古典的経済学 vs. 行動経済学

- 伝統的考え

- 合理的エージェント
- 合理的行為者 = **エコノ**
 - 完全情報
 - 期待効用に基づいた選択



- 行動主義的、心理学

- 人(**ヒト**)はときにはまったく合理的でなく、古典的な効用理論に基づいて行動しない



人間は理性的か？

- 論理問題をといていない
- 普通のひとは**不完全性定理**に悩まない



- 自己欺瞞
- 認知的不協和理論
- 誤帰属



参照点

- Bernoulli: “富のどんな小さな増加に起因する利得も、過去に保有していた財産量に反比例するはず”
- 冒険的は機会(賭け)の評価は、結果とし得られる財産よりも、可能な利得・損失の起こる参照点により依存する.
- “選好は、参照点の変化で操作される” (Tversky)
- しあわせの度合いは、当初の富の状態に対する**変化**によって決まる

参照点の例

ベルヌーイの理論:

どちらのOptionでも全く同じ選択肢である

∴ どちらのOptionでも同じようにえらぶはず

- Option 1: あなたは現在の富に加えて1000ドルもらったうえで、次のどちらかを選ぶように言われました。

- 50%の確率で1000ドルをもらう。
- または、確実に500ドルをもらう。

同じ確率で
1000ドルまたは2000ドル

確実に1500ドル

- Option 2: あなたは現在の富に加えて2000ドルもらったうえで、次のどちらかを選ぶように言われました。

- 50%の確率で1000ドルを失う。
- または、確実に500ドルを失う。

同じ確率で
1000ドルまたは2000ドル

確実に1500ドル

参照点の例

ところが、実際には。。。

- Option 1: あなたは現在の富に加えて1000ドルもらったうえで、次のどちらかを選ぶように言われました。

- 50%の確率で1000ドルをもらう。
- または、確実に500ドルをもらう。

多くの人が確実さを選ぶ

参照点は+1000ドル ∴ 500ドル増えるのがいい

- Option 2: あなたは現在の富に加えて2000ドルもらったうえで、次のどちらかを選ぶように言われました。

- 50%の確率で1000ドルを失う。
- または、確実に500ドルを失う。

多くの人がギャンブルを選ぶ

参照点は+2000ドル+ ∴ 500ドル確実に減るのはいやだ

不変性の喪失

- 不変性: A が B より好まれ、B が C より好まれるなら、A は C より好まれる
- チケットをもって映画に行くとき
 - \$40 のチケットをなくした
 - もう一枚チケットを買うか？
- 映画に行くので、チケットを映画館でかうとき
 - \$40 のチケット代をなくした
 - もう一度 \$40 を用意してチケットを買うか？
- 多くの人々は失ったチケットを再び買うことはしない(a)が、\$40を出してチケットを買う(b)ことはする

授かり効果 (The endowment effect)

- “Ownership creates satisfaction”
 - 所有感は、満足感を生む
- ヒトは自分のものになると、より価値をおく

Thaler, R. (University of Chicago),
Toward a positive theory of
consumer choice. Journal of
Economic Behavior and
Organization, March, 39-60, 1980.

学生が大学のマグカップをもらう。そのマグカップを売るならいくら出するかを答えてもらう。



マグカップをもらわなかった学生は、それを得るのにいくらはらうかを答えてもらう。



次のうちどうなったか？

- a) マグカップのある学生がより高く評価した.
- b) マグカップのない学生がより高く評価した.
- c) 両方の学生が同じように評価した.

マグカップのある学生の
平均販売金額

\$4.50



マグカップのない学生の
平均支払い金額

\$2.25



Kahneman, D. (UC Berkley), Knetsch, J. (Simon Fraser U), Thaler, R. (Cornell), 1990, Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy*, 98(6), 1325-1348.

Class A

はじめに学生にマグカップが与えられる。最後に、それをチョコレートバーと交換するか提案される。



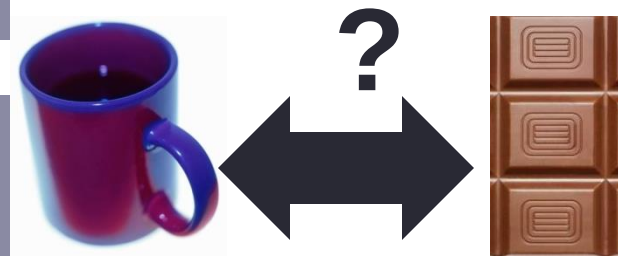
Class B

学生にチョコレートバーが与えられる。次にそれをマグカップと交換するか提案される。



Class C

はじめから、チョコレートバーかマグカップかどちらにするか提案される。



どのクラスがもっとも
マグカップを選びやすいか？

Class A

89%

マグカップを選ぶ



Class B

10%

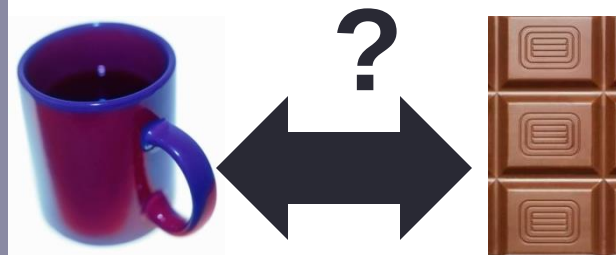
マグカップを選ぶ



Class C

59%

マグカップを選ぶ



J. Knetsch (Simon Fraser U.),
1989, The endowment effect and
evidence of nonreversible
indifference curves. *American
Economic Review*, 79, 1277-1284.

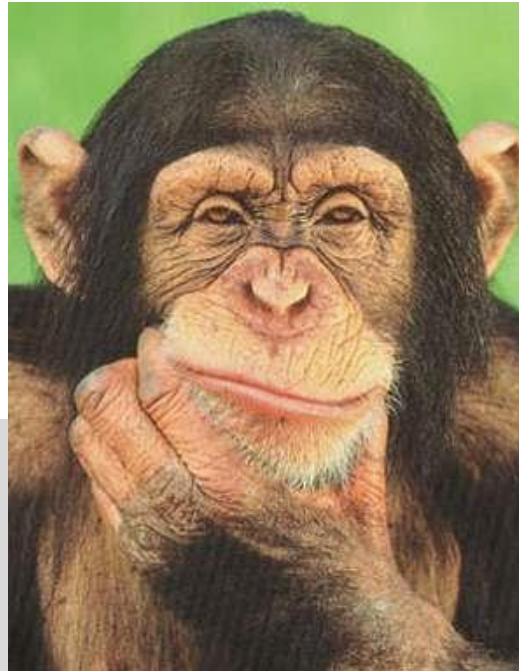
33匹のチンパンジーに、アイスクャンディーかピーナッツバターのいずれか(どちらも好物である)を与えて、もう一方と交換できるようにした。



ピーナッツバターが
先

89%

ピーナッツバターを
選ぶ



アイスクャンディーが
先

42%

アイスクャンディーを
選ぶ

Brosnan, S. (Emory), et al (Texas, Vanderbilt), 2007, Endowment effects in chimpanzees. *Current Biology*. 17. 1704-1707.

授かり効果とマーケティング



- Money-back guarantees : 返金保証
 - 商品を試すのをためらうことを減じる効果がある
- 一度でも商品を試してみると、授かり効果によって取引が成立する
- IKEA効果
 - 人はより多くの資源をつぎ込むと、その所有感覚が増加する
 - 自分が作ったものには愛着がわく
 - 認知的バイアス



The IKEA effect: When labor leads to love
Michael I. Nortona, Daniel Mochonb, Dan Arielyc,
Journal of Consumer Psychology,
Volume 22, Issue 3, July 2012, Pages 453–460

仮想的な所有感覚

- ある商品を所有したり、いっしょに暮らすと想像すると、あたかもすでに所有したような感覚になる
- このvirtual ownershipによって、購買や実際の所有に誘導される
- 成功する広告は商品と一緒に性格を生き生きと想像させることにたけている

損失回避

- “People are more motivated by avoiding a loss than acquiring a similar gain”
 - 人々は、同程度の利得を得るより、損失を避けることに動機づけられる
- フレーミング効果
 - 同じ選択が利得でなくて損失でフレームされると、異なる決定がなされる

フレーミング効果

提案1: 子供に対する控除額は、低所得者よりも高所得者を多くすべきですか？

⇒ NO! 金持ちに有利な税額控除なんてとんでもない

提案2: 子供のいない低所得者は、子供のいない高所得者と同額の追加納税をはらうべきですか？

⇒ NO! とんでもない

BUT よく考えてみよう。

論理的に考えてみれば、両方を拒絶することはできない。

損失回避

致命的な感染症に600人がかかるでしょう。
以下の2つの対策のどちらを選ぶか？



72% A: 400人が死ぬ

28% B: $\frac{1}{3}$ の確率で 600 人が助かり、 $\frac{2}{3}$ の確率で一人も助からない

損失回避

致命的な感染症に600人がかかるでしょう。
以下の2つの対策のどちらを選ぶか？



22% A: 200人が助かる

78% B: $\frac{1}{3}$ の確率で 誰も死なず、 $\frac{2}{3}$ の確率で600人が死ぬ

Tversky, A. & Kahneman, D., 1981,
The framing of decisions and the psychology
of choice. *Science*, 211, 453-458.

損失回避とフレーミング効果

600 人が死ぬ...

1/3 の確率で 誰も
死なず、2/3 の確
率で600人が死ぬ

=

1/3 の確率で 600
人が助かり、2/3
の確率で一人も助
からない

78%

≠

28%

損失を避けるのに多大のリスクを負う。
同じ選択を損失としてフレームし直すと決定が変更される。

以下の2つから選ぶ

A) \$240の利得

84%

B) 25% の確率で
\$1000の利得, 75%
の確率で何も得ら
れない

16%

余分の利
得を得るた
めにリスク
を負うこと
はすくない。

以下の2つから選ぶ

A) \$750の損失

13%

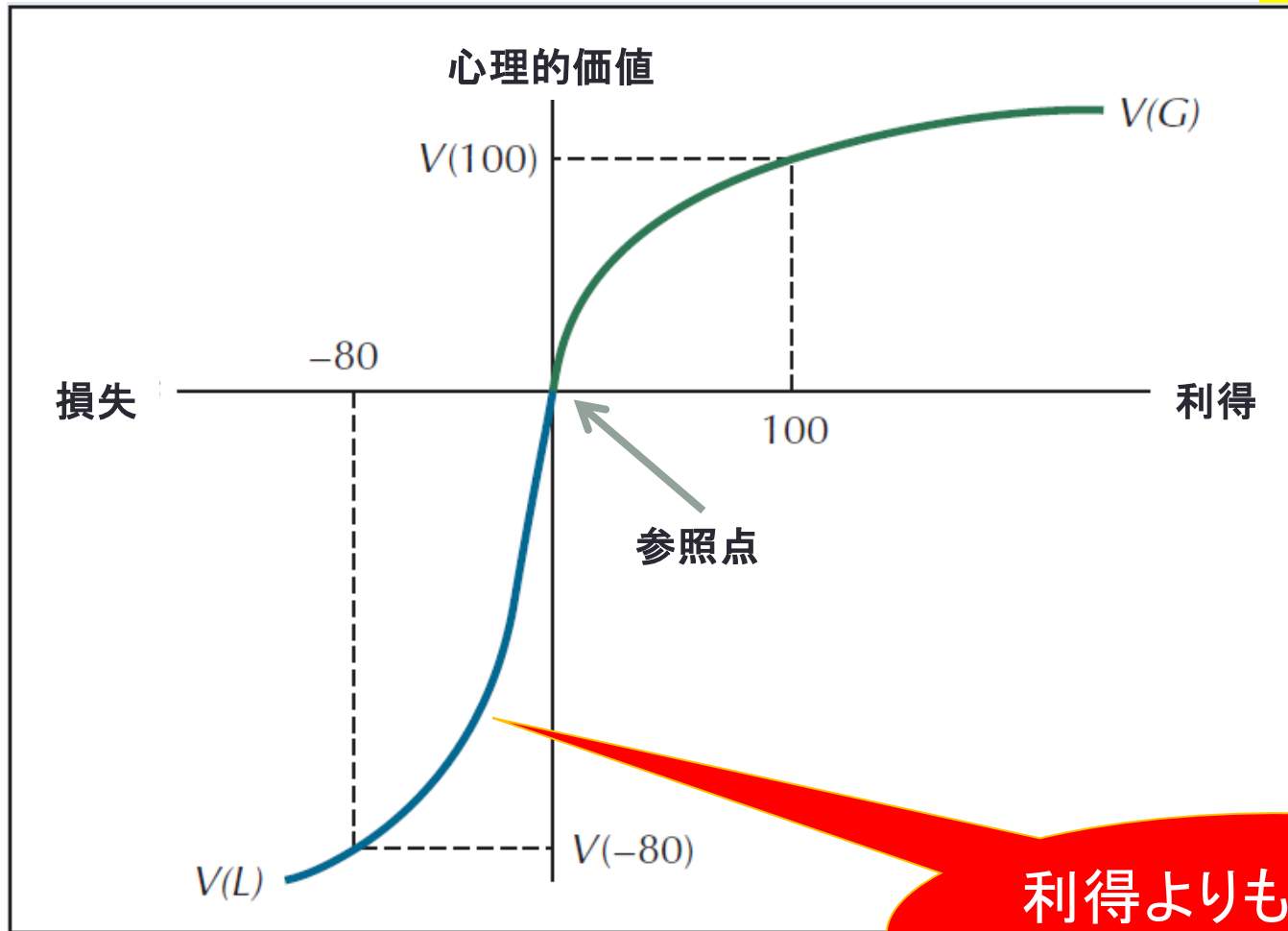
A) 25% の確率で
\$1000の損失, 75%
の確率で何も損しな
い

87%

余分の損
失を避ける
ためにリス
クを負いや
すい。

Kahneman-Tversky 価値関数

非対称な関数



利得よりも損失で
より急峻な傾きとなる

Kahneman-Tversky 価値関数

1. 人は利益と損失を非対称に扱う。
2. 意思決定において利得より損失に非常に重きを置く。
3. これは、必ずしも不合理なふるまいを意味ものではない。
4. 人は最初に項目を個別に評価し、個々の価値を合計する。



例：ゴルフのパッティングでバーディ／パーを狙う場合
250万回分のデータを分析
パー狙いはバーディ狙いよりも成功率が高い
これは、パットの難易度やカップからの距離と無関係

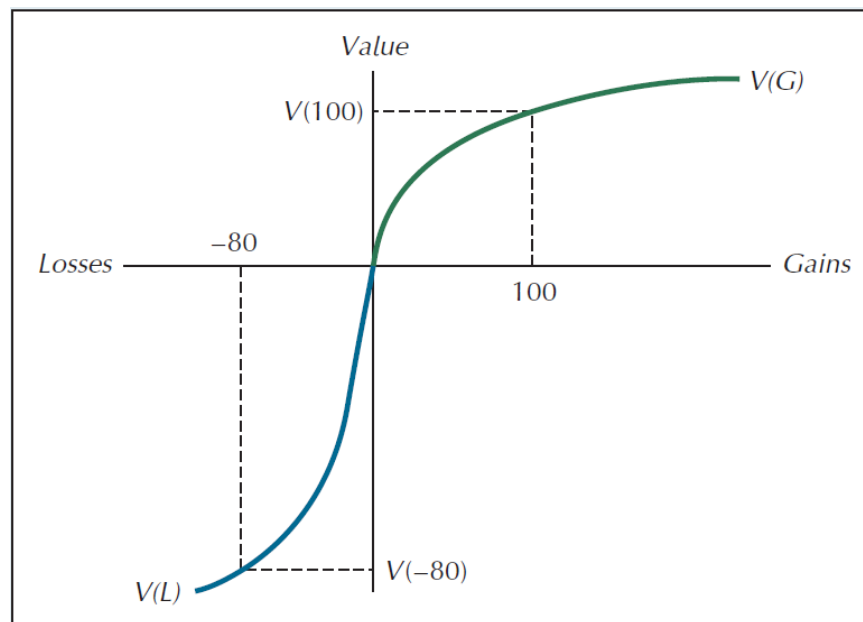
Kahneman-Tversky 価値関数

利得も損失もありうるギャンブルでは損失回避になり、極端に**リスク回避的**選択をとる

損失の傾き > 利得の傾き
(グラフの左) (グラフの右)



起こりうる損失が起こりうる利得の数倍も強く感じられる



Kahneman-Tversky 価値関数

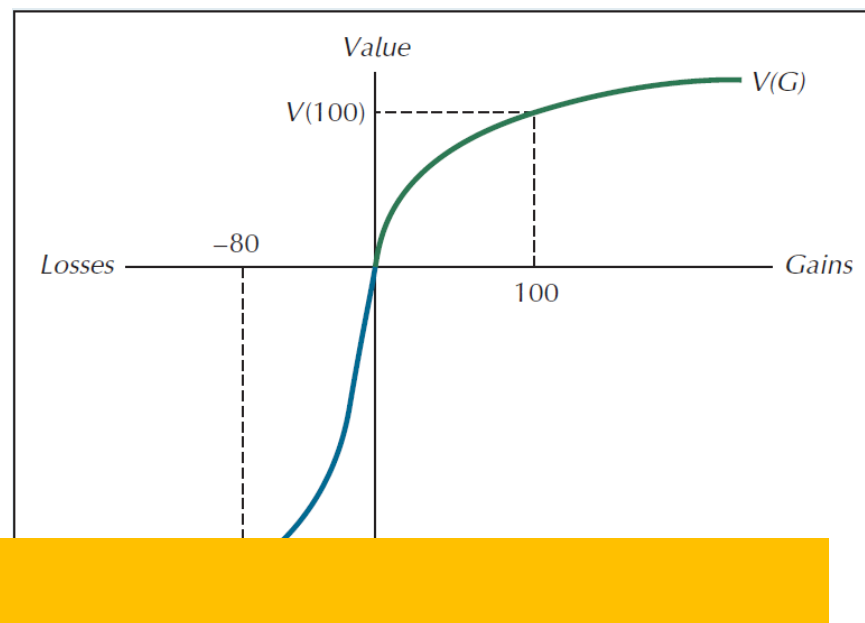
確実な損失と不確実だが大きな損失のように、どちらに転んでも悪い結果となるギャンブルでは、**リスク追及的**になる

損失領域のグラフの傾きは次第に緩やかになる

まとめ:

人間が意思決定するとき、

- * 損失の領域(左側)ではリスクを追及する
- * 利得の領域(右側)ではリスクを回避する



損失に対して、
よって、リ

効用理論の間違い

Risk Aversion and Expected-utility Theory: A Calibration Theorem, Matthew Rabin, *Econometrica*, Volume 68, Issue 5, pp.1281–1292, 2000. .

賭け1: 50%の確率で100ドルを失うが、50%の確率で200ドルをもらえる

賭け2: 50%の確率で200ドルを失うが、50%の確率で2万ドルをもらえる

多く人は、賭け1には手を出さない。

古典的効用理論に従うと、賭け1を断る人は、賭け2も断ることが証明できる。

そんなはずはない!!

正気の人間なら、賭け2を断ることはないだろう

Rabinの定理

50%でgドルを得て、50%で100ドルを失う賭けを断る人が、50%でLドルを失う賭けを断る金額の最大値

L	g			
	101	105	110	125
400	400	420	550	1,250
600	600	730	990	∞
800	800	1,050	2,090	∞
1,000	1,010	1,570	∞	∞
2,000	2,320	∞	∞	∞
4,000	5,750	∞	∞	∞
6,000	11,810	∞	∞	∞
8,000	34,940	∞	∞	∞
10,000	∞	∞	∞	∞
20,000	∞	∞	∞	∞

50% の確率で 100 ドルを失うが、50% の確率で 200 ドルをもらえる賭けを避ける人を考えます。この人の x ドルの効用関数を $U(x)$ と表します。このとき、この効用関数には、

Proof

- 単調増加である、つまり $U'(x) > 0$ となる
- 上に凸である、つまり $U''(x) < 0$ となる

の二つを仮定してよいでしょう。第 2 の性質は、効用関数が限界効用逓減になることを意味します (137 頁参照)。さて、先の人の所持金を w ドルとすると、賭けを避けるということから以下が導かれます。

$$U(w) > 0.5 \times U(w - 100) + 0.5 \times U(w + 200)$$

これは賭けをした後の自分の所持金の効用の期待値よりも、現時点での所持金の効用の方が大きいからです。 U が上に凸 (凹型) であることから $U(x)$ の傾きは x が大きくなるほど小さくなるので、

$$U(w + 100) > 0.5 \times U(w) + 0.5 \times U(w + 200)$$

です。この二式を足し合わせると、

$$\frac{1}{2}(U(w) - U(w - 100)) > U(w + 200) - U(w + 100)$$

が成り立ちます。これは任意の w で成り立つので、 w を $w - 100$ とすると、

$$\frac{1}{2}(U(w - 100) - U(w - 200)) > U(w + 100) - U(w)$$

再び、効用とは何か？



- とても痛い注射がある。決して慣れることはなく、毎日同じように痛い。
- このとき、患者にとっていかの2つは同じ効用だろうか？
 - 注射の回数を20回から18回に減らす
 - 注射の回数を6回から4回に減らす
- 古典的効用理論によると、
 - 注射の回数を20回から18回に減らす $\Rightarrow$ 1/10減
 - 注射の回数を6回から4回に減らす $\Rightarrow$ 1/3減 **よって、効用大**
- **本当か？ 何か変だ？**

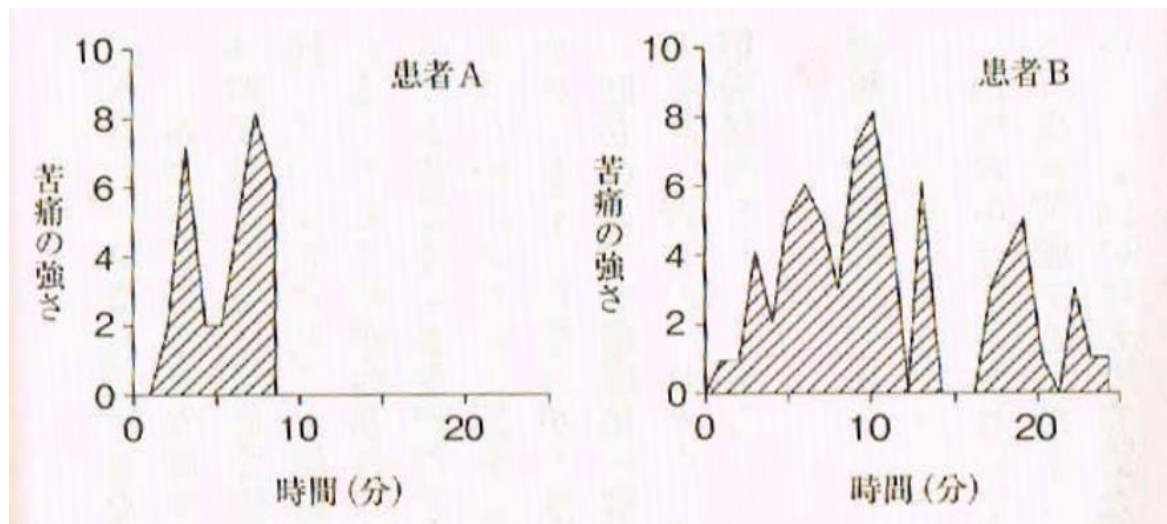
決定効用

- ピークエンドの法則

- 記憶に基づく評価は、ピーク時と終了時の苦痛・快樂の平均でほとんどきまる

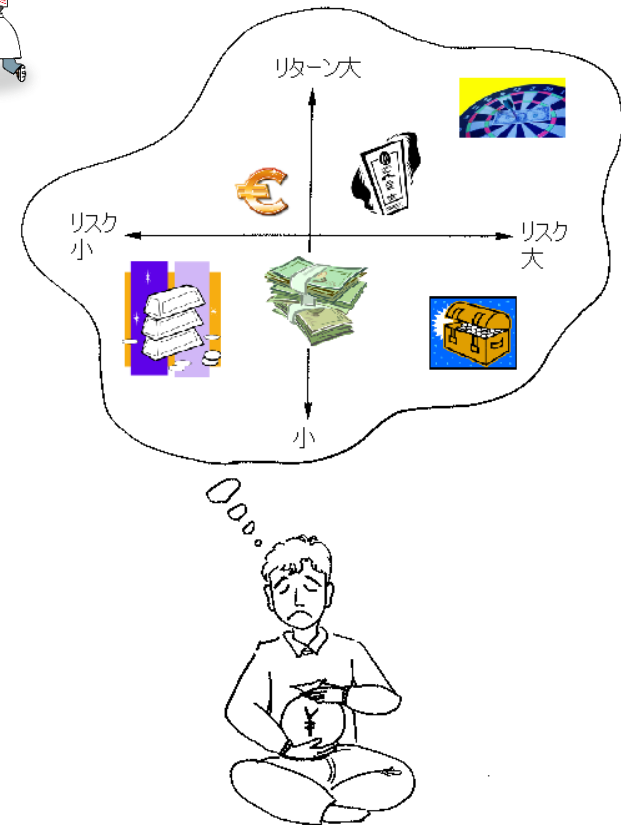
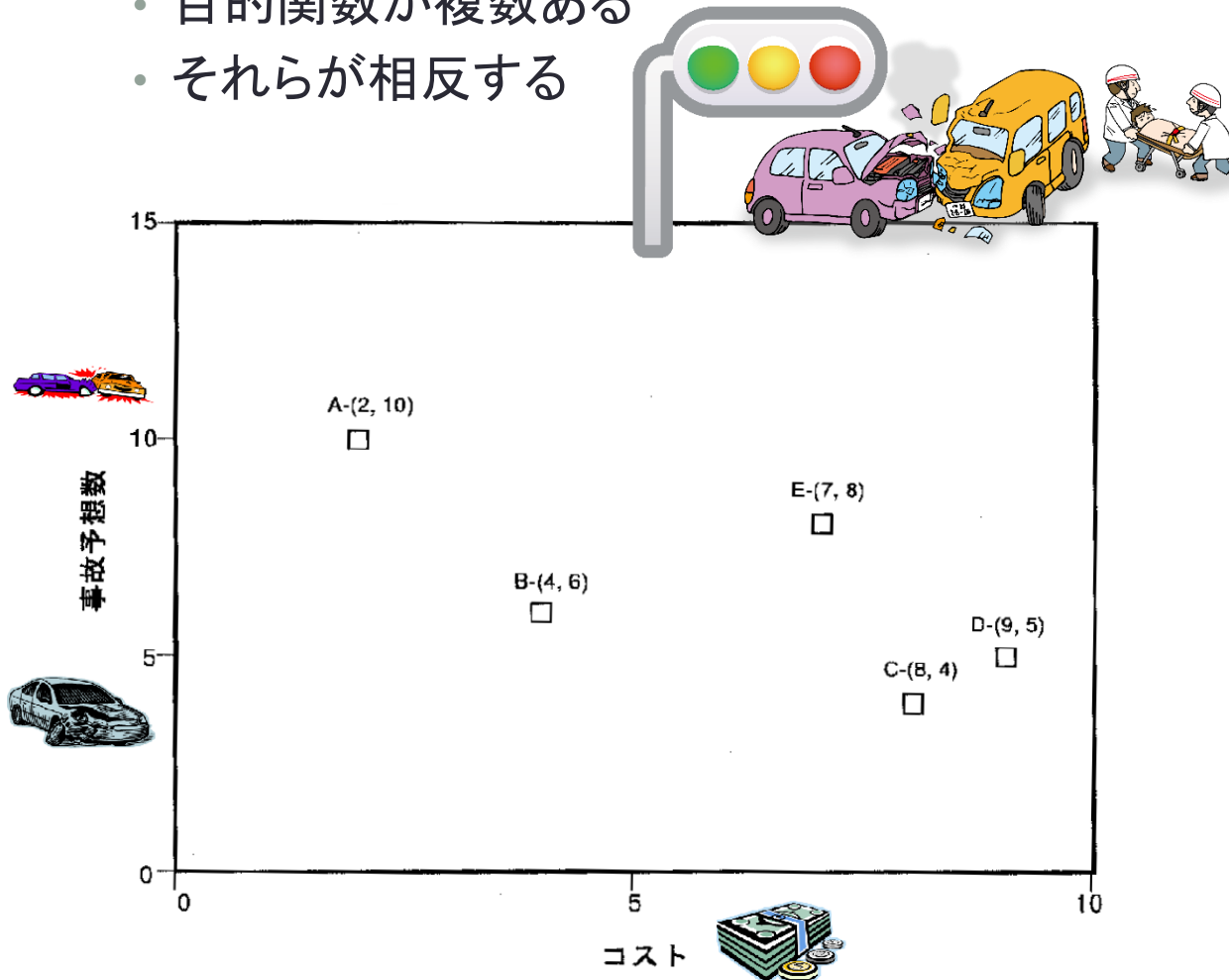
- 持続時間の無視

- 苦痛・快樂の持続時間は、総量の評価にはほとんど影響を及ぼさない

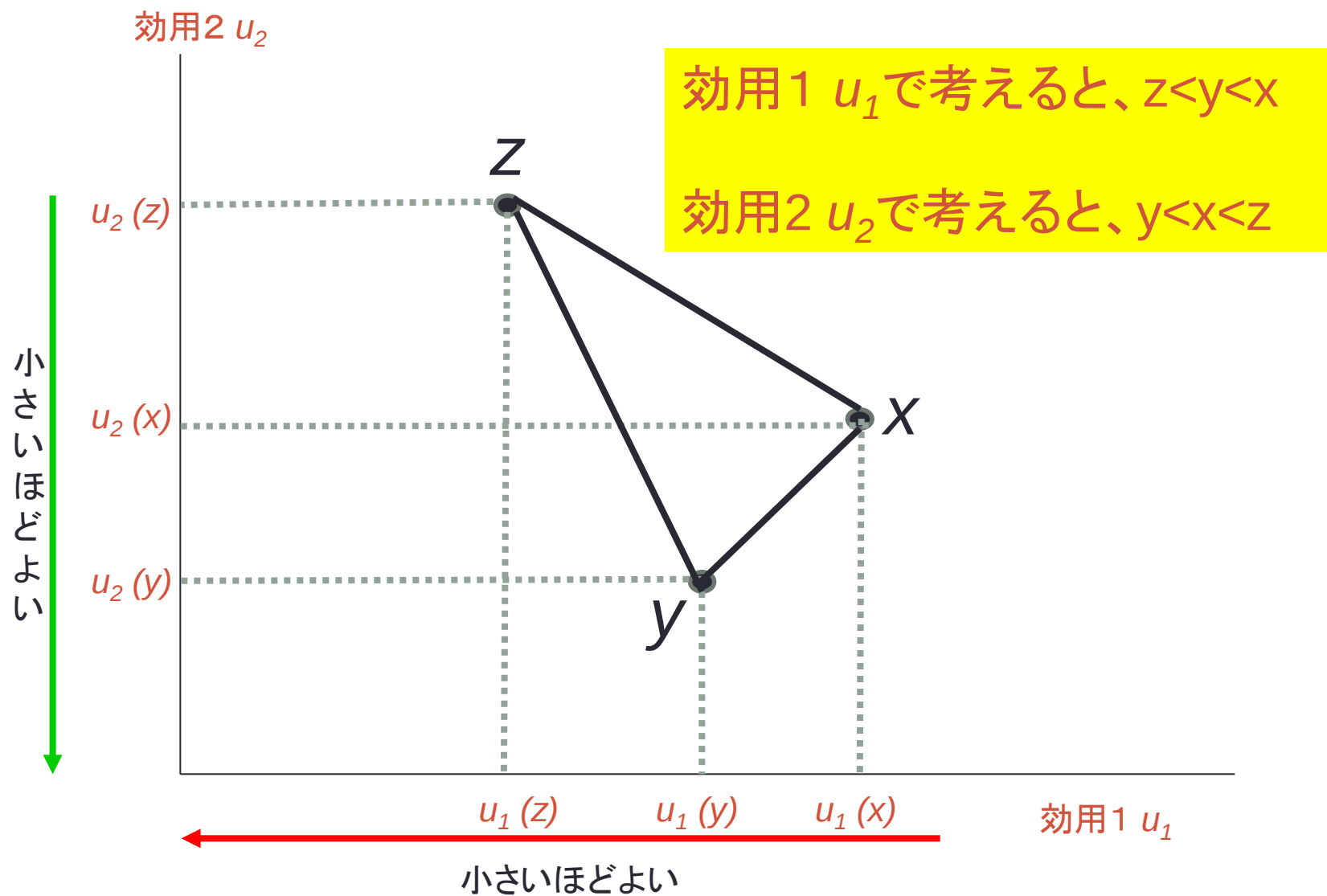


多目的最適化とは？

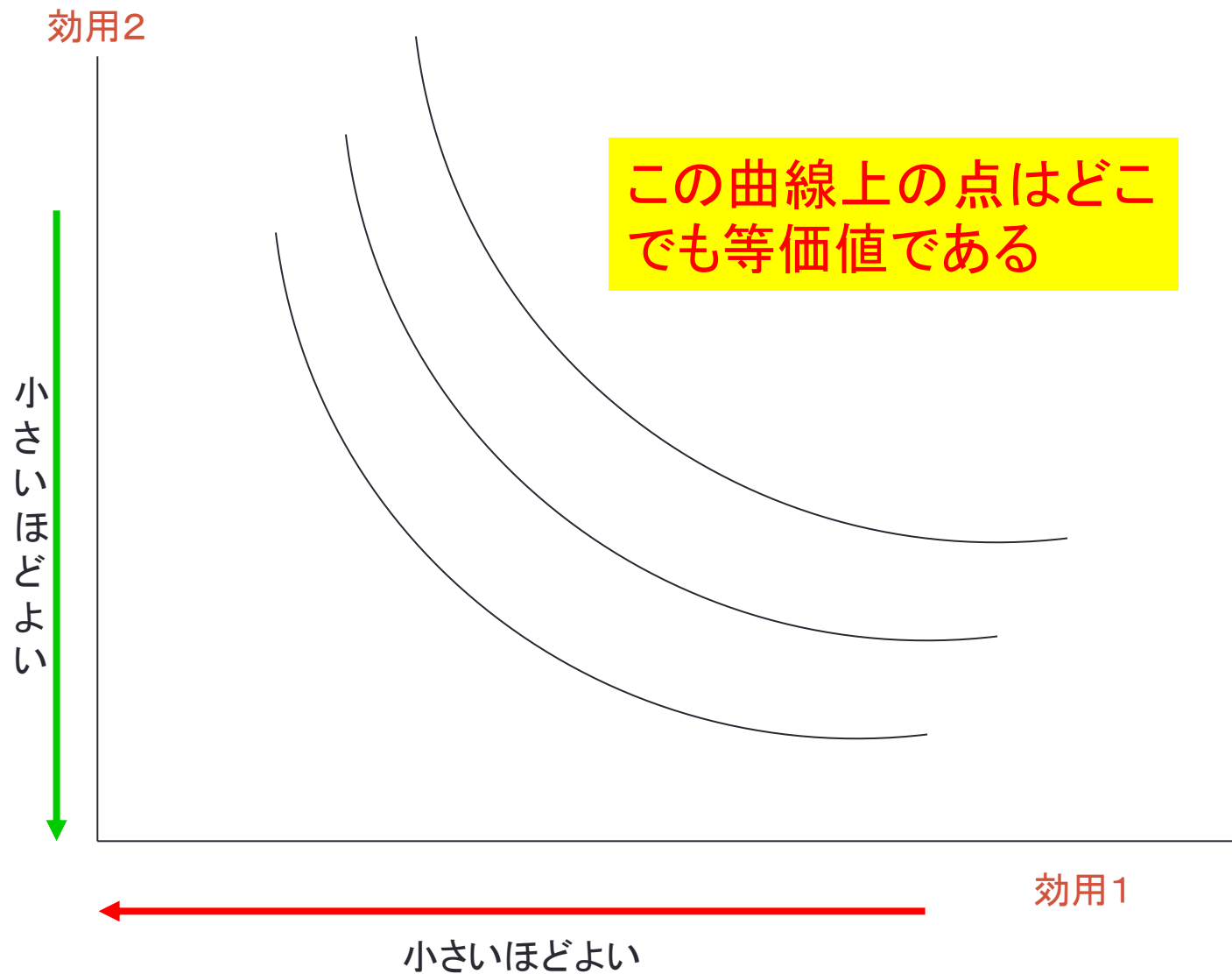
- 同時に複数の価値を最適化したい
 - 目的関数が複数ある
 - それらが相反する



多次元効用関数



無差別曲線



賭けにどう対処するか？

2-armed bandit

- 2本のレバーがある**スロットマシン**に使うためにN枚のコインが与えられた



Exploration (調査) $\mu_R > \mu_L$ なのか $\mu_R < \mu_L$ なのかを決定する.

Exploitation (利用) 上の決定にしたがって支払い率の良いレバーに賭ける.

- あなたは $m_1, m_2, \sigma_1, \sigma_2$ の値は知らない。
- 目標はN回の試行で得られる**賞金総額を最大化**することである。
- このとき各レバーへの試行配分(残りのコインの賭け方)はどのようにすればいいか？

2-armed bandit

Exploration (調査) $\mu_R > \mu_L$ なのか $\mu_R < \mu_L$ なのかを決定する.

Exploitation (利用) 上の決定にしたがって支払い率の良いレバーに賭ける.

- UCB値を用いる(ゲームの探索:UCTアルゴリズム):割と良い

$$\text{UCB 値} = \frac{\text{注目しているノードで勝った数}}{\text{注目しているノードに割いたプレイアウト数}} + c\sqrt{\frac{2\log(\text{全プレイアウト数})}{\text{注目しているノードに割いたプレイアウト数}}}$$

レバー*i*のこれまでの賞金

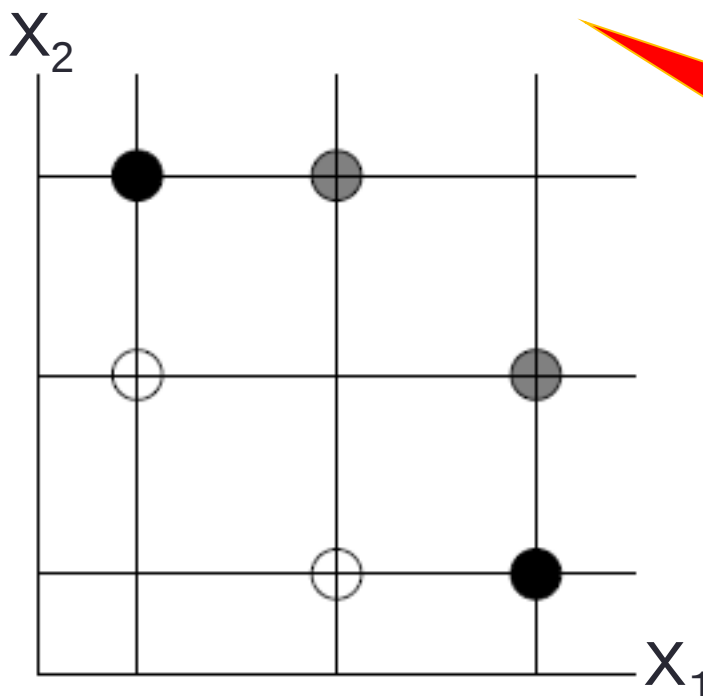


$$\text{UCB 値} = \frac{Q_n(i)}{n_i} + \sqrt{\frac{2\log n}{n_i}}$$

レバー*i*を引いた数

トムセン条件

$$X \succ Y \iff u_1(X) + u_2(X) > u_1(Y) + u_2(Y)$$

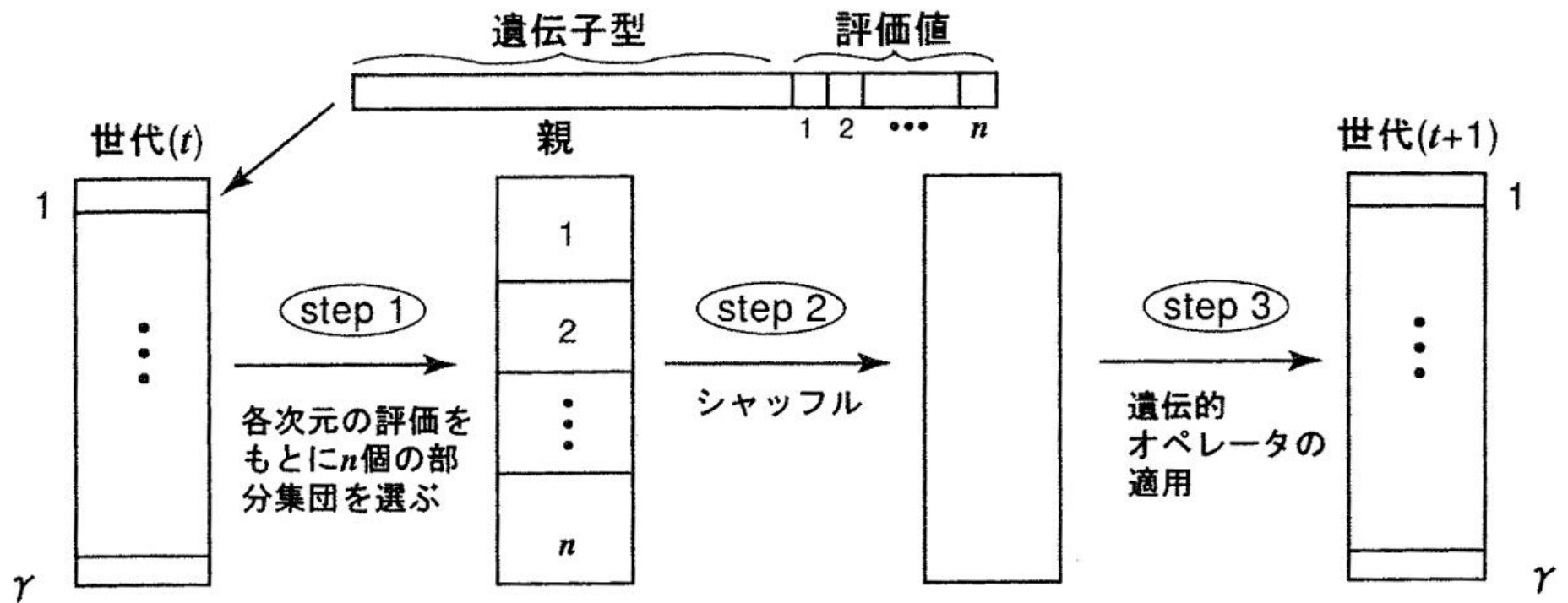


このようになるのが
理想的である

定義 5.5 トムセン条件

すべての $a, b, c \in X_1$ 、 $p, q, r \in X_2$ に対して、もしも $(b, p) \sim (a, q)$
かつ $(c, p) \sim (a, r)$ なら、 $(c, q) \sim (b, r)$ が成り立つ。

VEGA: GAによる多目的最適化



VEGA: シミュレータ紹介

