

行動選択における反応間間隔と 遅延時間の影響

期待報酬量が一定の状況での反応間間隔分布と
系列依存性および遅延価値割引判断について

川嶋健太郎

本論文では行動選択における時間の影響を検討するため、反応間間隔(interresponse time: IRT)と遅延時間という時間を取り上げて3つの研究を行った。報酬を獲得するまでの時間には遅延価値割引実験での遅延のように固定されていて被験者が変更できないものと、被験者が自由にその時間を決めることのできる一種の待ち時間に分類できると考えられた。本論文ではこの待ち時間として IRT を取り上げて、報酬を獲得するまでの時間(遅延および IRT)経過に伴って報酬の量や強化確率といった性質が変化することが被験者の反応や価値判断に与える影響が検討された。それぞれの研究では客観的基準から見るとどの時点での反応・判断であっても結果に違いがないように設計された実験条件を用意した。IRT について実験を行った研究1・2においては単位時間当たり報酬量を客観的基準とし、遅延価値割引実験を行った研究3においては名目利子率による指数関数的な割引を客観的基準とした。このような実験条件においても特定の時点での反応や報酬量の選択がされるならば、その時点・報酬量が主観的に見て最も高い価値を持つであろうと推定される。被験者による主観的な価値判断を反映した実験結果を説明するために、IRT の長さにより報酬量を割り引くモデルおよび経済的要因が割引率に影響を与えるとした遅延価値割引モデルを新たに提案した。

研究1「反応間間隔と報酬量についての実験心理学的研究」では IRT に応じて報酬量が増加する強化スケジュールを用いて、報酬量が IRT の長さを与える影響を2つの実験によって検討した。実験の結果、IRT の長さによって単位時間当たりの報酬量が異なる場合には、報酬量が最大となる IRT 付近での反応が見られるが、IRT の長さに関わらず単位時間当たりの報酬量が一定の場合には、強化確率によって生じられる IRT の長さが変わることが示された。

研究2「反応間間隔と強化確率、および系列依存性についての実験心理学的研究」では IRT に応じて強化確率が上昇する強化スケジュールを用いて、IRT 分布および IRT 系列依存性・IRT 変動性を3つの実験によって調べた。実験の結果、被験者によって頻度の高い IRT が異なること、IRT 系列依存性が認められることが示された。また IRT のばらつきである変動性も強化スケジュールによって操作可能であることが示された。

研究3「遅延価値割引における利子率・インフレ率の効果についての実験心理学的研究」ではシミュレーションゲーム的な実験課題を用い、遅延時間と報酬量の選択において、名目利子率・インフレ率という経済的要因がどう影響を与えるかについて4つの実験によって検討した。実験の結果、名目利子率が一定であるため客観的には同一の判断をするべきであっても、インフレ率によって遅延価値割引判断に違いが見られることが示された。またインフレ率が一定の状況での名目利子率の効果、および実質利子率が一定である場合での名目利子率・インフレ率の効果が見られた。

以下、構成に沿って本論文を要約する。

第1章 序論

ここでは報酬を獲得するまでの時間には遅延価値割引実験において一般的に用いられているような固定的で変更することのできない遅延と、自由に変更することのできる一種の待ち時間の 2 種類があるという本論文での基本的な見方を説明した。またこの遅延・待ち時間という時間経過中の報酬の性質(報酬量や強化確率など)が変化することが、人間の行動選択や価値判断に影響するのではないかと考えられた。

1.1 対応法則と最大化

ここでは本論文の 3 つの研究それぞれと関係の深い対応法則と最大化理論について概観した。対応法則では 2 つの行動の選択比率はその行動から得られる強化子の質や量の比率と対応すると考える。一方、巨視的最大化理論では動物・人間は強化率(単位時間当たり報酬量)が最大になるように行動を選択すると考える。対応法則と最大化のどちらが選択行動をうまく説明できるか盛んに議論されてきた。最後に対応法則と最大化理論をつなぐものとして、IRT が一種の遅延時間と考えられること、また IRT による割引という概念が提案されていることを示した。

1.2 反応間間隔について

反応間間隔(IRT)はフリーオペラント実験におけるもっと基本的な微視的行動指標である。ここでは IRT 強化理論や IRT の長さが強化によってコントロール可能であるとする実験結果などを紹介した。また連続した IRT の系列の中で、IRT 同士に見られる相関関係である IRT 系列依存性があることを支持する研究と、支持しない研究を紹介した。

1.3 遅延価値割引

遅延価値割引とは報酬が得られるまでの遅延によって、報酬の価値が低下することである。ここではまず双曲線型($V=A/(1+kD)$)・指数型割引関数($V=A \cdot \text{Exp}(-kD)$)をはじめとしたいくつかの割引関数、経済学における時間割引、規範的な指数型モデルからの数々の逸脱現象、遅延価値割引に影響を与えるその他の要因(年齢・収入など)や脳科学・機械による強化学習との関連について説明した。

第2章 研究 1：反応間間隔と報酬量についての実験心理学的研究

研究 1 では直線的報酬量増加(LRAI)、およびS字型報酬量増加(SRAI)スケジュールを提案し、どのような長さのIRTの頻度が高いか検討した。図 1はLRAIスケジュールでの、図 2はSRAIスケジュールでのIRTの長さに対する 1 回あたりの報酬量と、1 試行あたりの報酬量を表している。LRAIスケジュールではIRTの長さに応じて直線的に反応 1 回あたりの報酬量が増加していくが、1 試行あたりの報酬量(および単位時間当たり報酬量)はどのようなIRTであっても同一である。一方、SRAIスケジュールではIRTに応じて反応 1 回あたりの報酬量がS字型に増加し、1 試行あたりの報

酬量は特定のIRTにおいて最大値をとる。

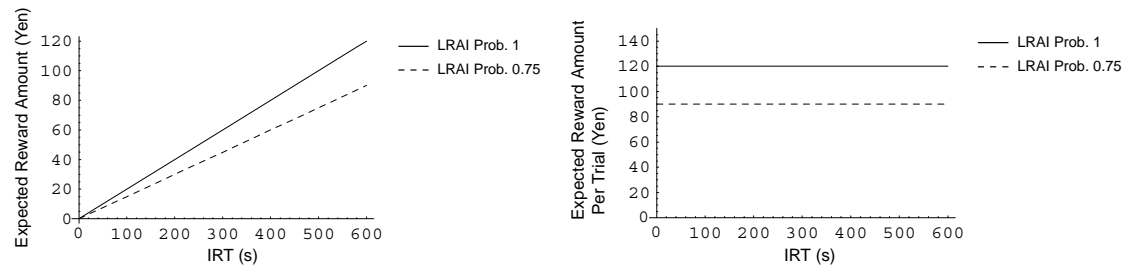


図 1 LRAI スケジュールでの反応 1 回あたりの報酬量(左)と 1 試行あたりの報酬量(右)

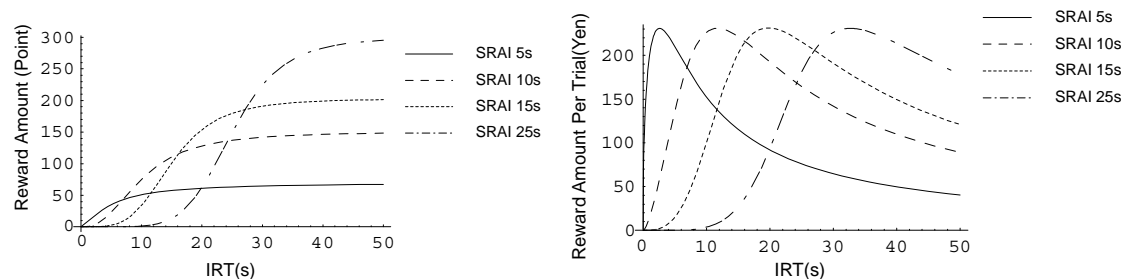


図 2 SRAI スケジュールでの反応 1 回あたりの報酬量(左)と 1 試行あたりの報酬量(右)

2.1 実験 1：S字型報酬量増加(SRAI)スケジュールでの実験

実験 1 では単位時間当たりの報酬量を最大化する IRT において反応の頻度が高くなるかどうかを調べるため、SRAI と LRAI スケジュールの並立スケジュールを用いた実験を行った。SRAI スケジュールでは特定の長さの IRT において単位時間当たりの報酬量が最大となる。実験の結果、SRAI スケジュールに対してすべての被験者が単位時間当たりの報酬量が最大となる長さ付近の IRT で反応をすることが示された。

2.2 実験 2：直線型報酬量増加(LRAI)スケジュールでの強化確率の影響についての実験

実験 2 では IRT の長さを調整することでは単位時間当たり報酬量を最大化することが出来ない LRAI スケジュールを用いて、強化確率の違いが IRT に与える影響を調べた。この結果、強化確率が 1 の場合にはしばしば試行時間とほぼ同じほどの長い IRT が観察される一方で、強化確率 0.75 の場合には 2s から 4s 程度の短めの IRT が多く観察された。強化確率が変化しても単位時間当たり報酬量は IRT の長さに影響されないため、実験 2 での被験者は単位時間当たり報酬量を最大化するために IRT の長さを変更したのだとはいえなかった。これらの結果は対応法則・巨視

的最大化理論などでは説明することが出来なかった。

2.3 総合考察

実験 1 から被験者は単位時間当たりの報酬量をほぼ最大化する IRT で反応することが示されたが、その一方で実験 2 の結果から IRT の長さを変えることでは最大化をすることのできない LRAI スケジュールにおいては、IRT が強化確率によって影響されることが示唆された。このことを説明するため IRT を一種の遅延・待ち時間と仮定して、IRT の長さによって報酬量を割引く 3 つのモデル(IRT 割引逆数・指数関数型・双曲線型モデル)を提案した。次の式は行動が 2 つの場合の IRT 割引指数関数型モデルである。

$$\begin{aligned} \mathbf{t}^* &\in \arg \max_{\mathbf{t}} \{V_{\text{exp1}}(t_1) + V_{\text{exp2}}(t_2)\} \\ &= \arg \max_{\mathbf{t}} [e^{-k_1 \cdot t_1} \{p_1(t_1) \cdot A_1(t_1) - C\} + e^{-k_2 \cdot t_2} \{p_2(t_2) \cdot A_2(t_2) - C\}] \end{aligned}$$

$$\text{Subject to } b \leq t_1 \cdot t_2, \quad t_l \leq t_1 \leq t_h, \quad t_l \leq t_2 \leq t_h$$

(1)

ここで t_1, t_2 はそれぞれ行動 1, 行動 2 での IRT の長さ, \mathbf{t}^* は最適な IRT の組(t_1^*, t_2^*), t_l, t_h はそれぞれ IRT の物理的な下限と上限, b は行動間の競合の程度を表すパラメータである。 $V_{\text{exp1}}, V_{\text{exp2}}$ はそれぞれ行動 1, 行動 2 の主観的価値, $p_1(t_1), p_2(t_2)$ はそれぞれの IRT に対する強化確率, $A_1(t_1), A_2(t_2)$ は IRT に対する報酬量, C は反応 1 回ごとのコスト, k_1, k_2 は強化確率によって変化する割引率である。このモデルでは $V_{\text{exp1}} + V_{\text{exp2}}$ を最大とするような \mathbf{t}^* が選ばれようとする。

実験 1 の SRAI スケジュールでの結果については 3 つのモデルのどれでも特定の IRT の頻度が高いことを示すことが出来た。次に実験 2 の LRAI スケジュールの結果については IRT 指数関数型モデルでのみ強化確率によって最適な IRT が変化しうることが示された。また IRT 指数関数型割引から単一強化スケジュールについての対応法則とほぼ同じ強化率と反応率の関係を導き出せることが示された。

第3章 研究 2：反応間隔の系列依存性についての実験心理学的研究

研究 2 では staircaseSRW および vSRW スケジュールを提案して、IRT 分布の測定、IRT 系列依存性とその個人差としての反応パターンの分析、および IRT 変動性を高めることを目的とした強化スケジュールを検討する実験を行った。図 3(左)は staircaseSRW と Random Ratio(RR) スケジュールでの IRT の長さに対する反応 1 回あたりの強化確率を、図 3(右)は同じく 1 試行あたりの期待報酬量をあらわしている。図のように staircaseSRW スケジュールでは IRT の長さに応じて強化確率を階段状(各 IRT 区間での強化確率は一定)に増加させている。こうすることで単位時間当たり報

酬量をほぼ一定に保つことが出来る。

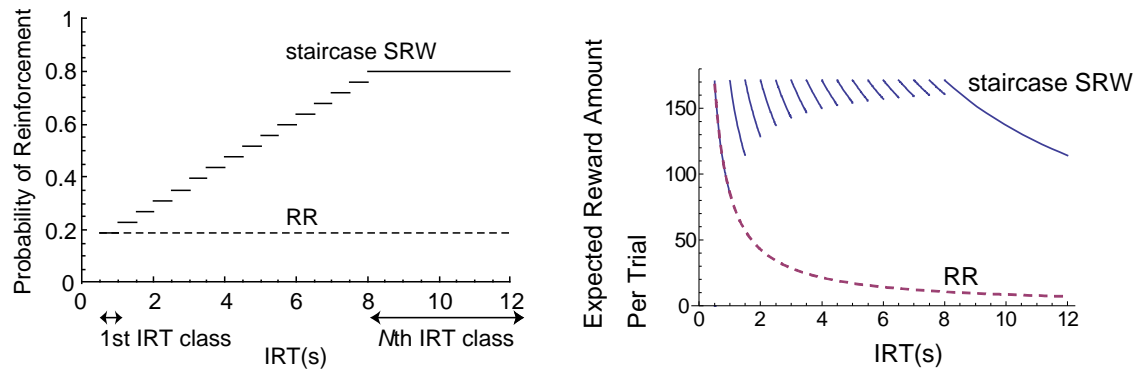


図 3 staircaseSRW スケジュールと RR スケジュールでの IRT に対する強化確率(左)と 1 試行での期待報酬量(右)

3.1 実験 3: 強化確率により単位時間報酬量を一定にしたスケジュールでの IRT 分布と IRT 系列依存性

実験 3 では staircaseSRW または RR スケジュールを用いて, IRT の系列依存性を調べた. この IRT 系列の自己相関関数の結果から, 多くの実験試行において IRT 系列は独立かつ定常な系列ではなく, 系列依存性が認められた. 図 4 は IRT_n と強化／反応コストに後続する IRT_{n+1} の相対頻度を表している. 図のように強化後・反応コスト後どちらも対角線上での IRT の頻度が高かった. このことはある IRT での反応の後にはほぼ同じ長さの IRT が続きやすいことを示しており, IRT には系列依存性が認められる.

3.2 実験 4: IRT 系列依存性の学習期間による変化についての実験

IRT 系列依存性の学習による変化を調べるため, 実験 3 と同じ手続きで実験期間を長くして実験を行った. この結果, 学習の初期と安定期では系列依存性に違いが見られた. IRT_n と IRT_{n+1} の相対頻度では最初の 2 日間は明瞭な系列依存性を示していたが, 最後の 2 日間のデータでは, 観察される IRT の範囲が狭まり, IRT 系列依存性は最初の 2 日間に比べて弱くなっていた. また学習の安定期では初期と比較して, 特定の IRT での反応が多くみられた.

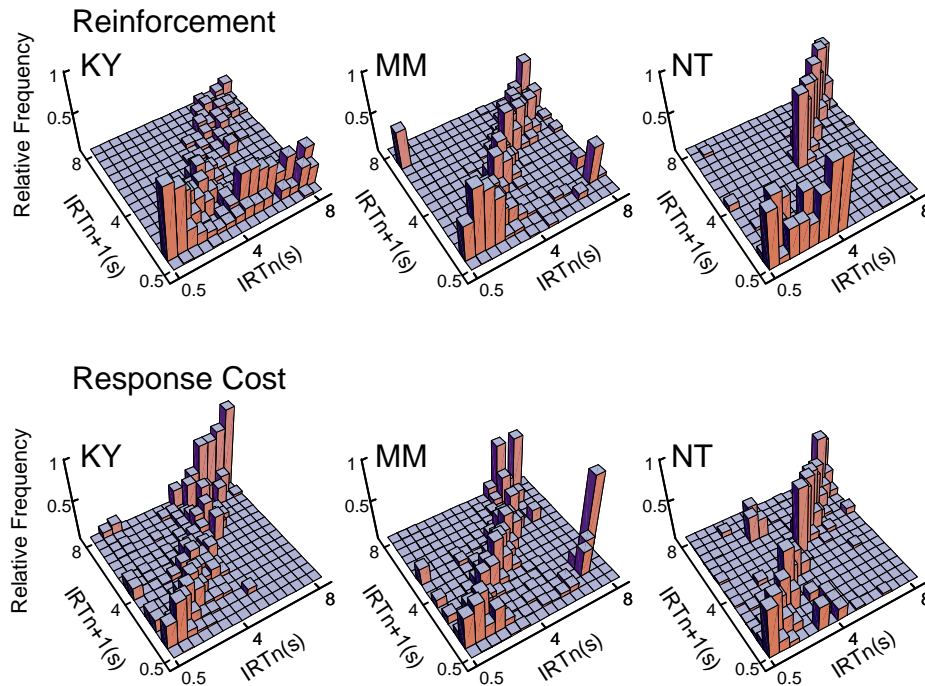


図 4 実験 3 での staircaseSRW スケジュールにおける IRT の系列依存性. IRT_n と強化に後続する IRT_{n+1} の相対頻度(上段). IRT_n と反応コストに後続する IRT_{n+1} の相対頻度(下段).

3.3 実験 5 : vSRW スケジュール下での IRT 変動性

vSRW スケジュールは, staircaseSRW スケジュールをもとに被験者による反応があった IRT クラスの強化確率を下げ, 他の IRT クラスの強化確率をその分上昇させることで単位時間あたりの報酬量をほぼ一定に保つ強化スケジュールである. staircaseSRW-vSRW-staircaseSRW という ABA デザインで実験を行った結果, vSRW スケジュールでは IRT 系列依存性は認められるものの, staircaseSRW スケジュールでの実験試行よりもばらつきで見た IRT 変動性が増加することが示された.

3.4 総合考察

実験 3・4 での staircaseSRW スケジュールでの結果から, 強化確率によって単位時間当たり報酬量を一定にした場合でも特定の長さの IRT が高い頻度で生起すると考えられた. そこで研究 1 で提案した IRT 割引双曲線型・指数関数型の 2 つのモデルを検討したところ, IRT 割引指数関数型モデルでのみ, 割引率を調整することで staircaseSRW スケジュールのもとで特定の IRT の頻度が高いことを説明できた. また IRT_n と IRT_{n+1} の相対頻度グラフから強化や反応コストに対して個人差があることが示唆された.

第4章 研究3：遅延価値割引における利子率・インフレ率の効果についての実験心理学的研究

研究3では固定された遅延報酬であっても、直後報酬の量の変化などによる影響を受けるか検討した。仮想的な経済における名目利子率により直後報酬の報酬量の増加を、インフレ率により直後・遅延双方の報酬の実質的な価値(購買力)をコントロールした。実験者が制御できないためにこれまでの遅延価値割引研究では利子率・インフレ率はほとんど取上げられてこなかったが、経済的な意味での割引においては最も重要な要因であるといえる。本研究では仮想通貨(モク)を用いたゲームの実験課題を使用することでこれらの経済的要因が遅延価値割引に与える影響を検討した。

実験課題はゲームステージ1(利子と商品価格変化を学習)ーボーナスステージ(遅延・直後報酬の選択)ーゲームステージ2(ボーナスステージでの選択どおりに遅延・直後報酬を受け取る)という3段階で構成されていた。図5左は実験課題でのゲームステージの概観を示している。ここでは被験者はターンという時間的区切りごとに、定期的収入および預金額からの金利収入を獲得し、商品の値段変化を観察した。またマウスをクリックすることで自由に商品を購入することが出来た。実験者はターンごとの商品の値段変化でインフレ率を、預金額への金利支払いで名目利子率をコントロールすることが出来た。図5右はボーナスステージの概観を示している。直後報酬と遅延報酬が表示され、被験者はマウスを用いて選択をした。被験者の選択によって呈示される直後報酬の金額が調整された。ここで選択をしたとおりに、後に続くゲームステージ2において直後報酬・遅延報酬が支払われた。

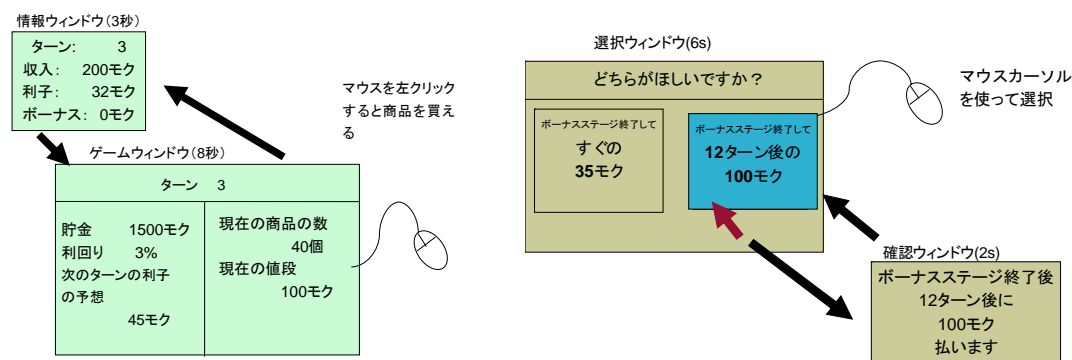


図5 実験課題でのゲームステージ(左)とボーナスステージ(右)の概観

例として表1は実験6で用いられた3つの条件(インフレ条件・ゼロインフレ条件・デフレ条件)での36ターン後の遅延報酬と複利の利子の付く直後報酬の購買力を比較したものを示している。すべての条件で名目利子率は共通で1%であり、実質利子率(=名目利子率-インフレ率)は0以上である。また商品価格は直後(0ターン)では100モクであり経済ごとのインフレ率の違いによって価格は変化していく。購買力(金額/商品価格)とは直後報酬・遅延報酬によって商品を何個分購入

可能であることを示している。遅延報酬は 36 ターン後の 100 モクであり、購買力はインフレ率によって異なっている。

研究 3 では遅延報酬の客観的な割引現在価値を「直後報酬＋複利の利子」の購買力が遅延報酬の購買力と等しくなるような直後報酬の金額であると考え、なぜならもしも直後報酬の購買力が遅延報酬の購買力よりも高いのならば直後報酬が選択され、逆ならば遅延報酬が選択されるべきだからである。被験者が購買力という観点から選択するならば、遅延報酬の遅延期間後の直後報酬＋複利の利子が遅延報酬と等しくなるまで呈示される直後報酬の金額が調整されることになる。実験 6 ではすべての条件で名目利子率は 1%であったため、すべての条件で客観的な割引価値は 69.89 モクとなった。

4.1 実験 6：インフレ率の効果についての実験

名目利子率を一定にして、インフレ率が遅延価値割引に影響を調べた。名目利子率が一定であるため客観的な割引価値はインフレ率による影響がないはずである。しかし実験結果からインフレ率により主観的な割引率に違いがあることが示された。

4.2 実験 7：利子率の効果についての実験 デフレ状況

デフレ状況での名目利子率が遅延価値割引に与える影響が調べられた。この結果、名目利子率が遅延価値割引に強い影響を与えることが示された。

表 1 実質利子率がマイナスでない場合での直後報酬(複利の利子がつく)と遅延報酬の購買力

ターン	商品価格	複利の利子のつく直後報酬		遅延報酬	
		金額	購買力	金額	購買力
インフレ条件(インフレ率:1%, 名目利子率:1%)					
0	\$100.00	\$69.89	0.699	－	－
12	\$112.68	\$78.76	0.699	－	－
24	\$126.97	\$88.74	0.699	－	－
36	\$143.08	\$100.00	0.699	\$100	0.699
ゼロインフレ条件(インフレ率:0%, 名目利子率:1%)					
0	\$100.00	\$69.89	0.699	－	－
12	\$100.00	\$78.76	0.788	－	－
24	\$100.00	\$88.74	0.887	－	－
36	\$100.00	\$100.00	1	\$100	1
デフレ条件(インフレ率:-1%, 名目利子率:1%)					
0	\$100.00	\$69.89	0.699	－	－
12	\$88.64	\$78.76	0.889	－	－
24	\$78.57	\$88.74	1.13	－	－
36	\$69.64	\$100.00	1.436	\$100	1.436

4.3 実験 8： 利子率の効果についての実験 インフレ状況

インフレ状況での名目利子率が遅延価値割引に与える影響が調べられた。この結果、インフレ状況では名目利子率の影響は見られなかった。インフレ状況では遅延報酬の客観的価値が名目利子率のみではなくインフレ率からも影響を受けるからであると考えられた。

4.4 実験 9： 実質利子率一定の実験

実質利子率とは名目利子率からインフレ率をひいたものである。実質利子率が同一だが名目利子率とインフレ率が異なる組み合わせを実験条件として実験を行った。実質利子率が等しいと名目利子率・インフレ率が異なっても実験セッション中に購入できる商品の数は等しくなる。実験の結果、実質利子率が等しくてもインフレ率・名目利子率の違いにより遅延価値割引判断に違いが出る事が示された。

4.5 総合考察

利子率・インフレ率の効果を説明するために、Linear Expectation モデルと Regression モデルを提案し、実験 6 から 9 の結果をどの程度説明できるか検討した。Linear Expectation モデルは利子率やインフレ率により直線的に金額が増加すると被験者が予想していると仮定したモデルである。

$$V = \frac{A_D}{(1 + u' \cdot I_1 D)(1 + v' \cdot I_2 D)} \quad (2)$$

ここで V は遅延報酬 A_D の主観的価値、 I_1 は実験条件のインフレ率、 I_2 は名目利子率、 u' はインフレ率に対するパラメータ、 v' は名目利子率に対するパラメータは D は遅延である。

Regression モデルは重回帰式により利子率・インフレ率から割引率 k を推定するモデルである。

$$V = \frac{A_D}{1 + (a + b_1 I_1 + b_2 I_2) D} \quad (3)$$

図 6 は双曲線型割引関数と Linear Expectation モデル、Regression モデルを実験 6 から 9 までの結果に当てはめたものである。ただし、双曲線型割引関数に対しては条件ごとの当てはめであり、他の 2 つのモデルについては各実験の 3 つの実験条件の結果について同時に当てはめを行い、得られたパラメータと実験条件のインフレ率・名目利子率から曲線を描いている。2 つのモデルはどちらも双曲線型割引関数をベースに考案したものであるため、実験結果への当てはまりの程度はほとんどの場合、双曲線型割引モデルとほぼ同じであった。しかし、Linear Expectation モデ

ルでは非線形回帰分析ができないこともあり, Regressionモデルのほうが当てはめやすいと評価した。

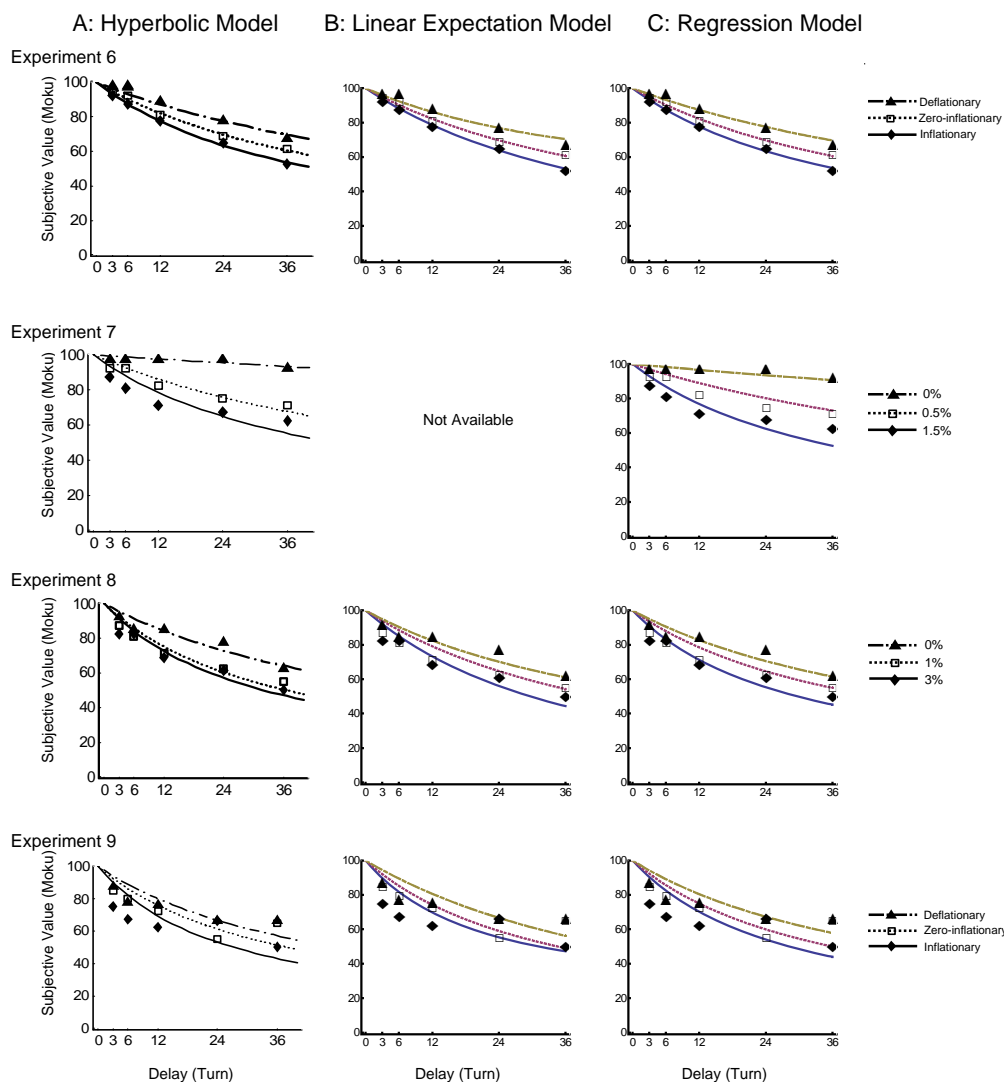


図 6 A:実験 6 から 9 での各実験条件の主観的等価点と双曲線型割引モデルと, B:Linear Expectation モデル, および C:Regression モデルへ 3 つの実験条件を同時に曲線あてはめした結果

第5章 総論

ここではまず研究 1 から 3 の実験結果をまとめた後, 研究 1 と 3 における特徴でもあった, 時間に応じた報酬量の増加と時間の価値について考察した. 研究 1 での LRAI スケジュールでは直線的に報酬量が増加するが, これは双曲線型割引と関係が深い. これらのことから IRT と報酬後の遅延の類似点と相違点について考察した.