

一般発表 | Track 5: 高度なデータ利活用・ドメイン応用（医療情報，教育，地理情報等）

2026年3月2日(月) 9:30 ~ 11:40 | K会場

[7K] アンケート分析

座長: 莊司 慶行(静岡大学) コメンテータ: 神門 典子(総合研究大学院大学)

9:30 ~ 9:55

[7K-01] 出産意欲についてのアンケート分析

*三原 裕子¹、丸山 亜希子¹、上田 真由美² (1. 流通科学大学、2. 追手門学院大学)

9:55 ~ 10:20

[7K-02] 就職活動における情報探索行動と時間選好の関係分析

*後藤 侑美¹、山本 修平¹ (1. 筑波大学)

10:20 ~ 10:40

[7K-03] 公開研修参加者が変化を実感する要因とは：満足度調査からの示唆

*永瀬 純¹、野口 怜¹ (1. 東京理科大学)

10:40 ~ 11:05

[7K-04] 段階評価データの分布形状差に着目した非類似度指標とウェルビーイングデータ分析への応用

*萩田 碧偉¹、山崎 綾一郎²、山岸 祐己^{1,3}、青木 成樹⁴、橋本 正洋^{4,5,6} (1. 静岡理工科大学、2. 静岡理工科大学大学院、3. 浜松医科大学、4. マリンオープンイノベーション機構、5. 法政大学、6. 東京科学大学)

11:05 ~ 11:30

[7K-05] 意思決定を特徴づける形容詞的修飾を用いた大規模言語モデルの実証的分析

*馬田 光琉¹、山本 修平¹ (1. 筑波大学)

出産意欲についてのアンケート分析

三原裕子[†] 丸山亜希子[†] 上田真由美[‡]

[†] 流通科学大学経済学部 〒651-2188 神戸市西区学園西町3丁目1番

[‡] 追手門学院大学経営学部 〒567-8620 大阪府茨木市太田東芝町1番1号

E-mail: [†] {Yuko_Mihara, Akiko_Maruyama}@red.umds.ac.jp, [‡] may-ueda@otemon.ac.jp

あらまし 現在日本では、これまでに例を見ない速度で少子化が進行している。夫婦の最終的な出生数である完結出生児数が、長期的に人口が増えも減りもせず一定規模に維持するために必要な人口置換率を下回っており、人口規模の縮小は避けられない状況にある。少子化の背景として、所得水準の伸び悩みや経済的不安定性、子育てに要する費用の増大、晩婚化の進行など、さまざまな要因が指摘されてきた。こうした中、45～49歳において1人も子どもを産んでいない女性の割合である生涯無子率はOECD加盟国の中で最も高く、今後もさらに上昇すると見込まれている。無子個人と有子個人をそれぞれ対象に、出産の意向について被説明変数として二項ロジスティック回帰分析を行った。その結果、無子個人については、年齢や経験したことのない出産に対する不安が出産意欲に大きく影響を及ぼすことが分かった。一方で、有子個人については、子育てに関連する費用が現実のものとして可視化されるため、子育てにおいて発生する負担などが強く影響していることが確認された。

キーワード 出産意欲、無子、アンケート分析

1. はじめに

日本では少子化が急速に進行している。厚生労働省[11]によると、令和6年の出生数は68万6061人で、これは前年の72万7288人より約4万人減であることが示される。また、母の年齢(5歳階級)別にみると、全ての階級で減少し、合計特殊出生率は1.15と前年の1.20より低下したことも合わせて示された。

夫婦の最終的な出生数である完結出生児数については、国立社会保障・人口問題研究所[13]によると、2021年(第16回)調査時点では1.90人となり、これは2002年(第12回)調査までは、2.2人前後で安定的に推移してきたのに対して減少に転じた。人口が長期的に増えも減りもせず、一定となるために必要な合計特殊出生率を人口置換率という。この人口置換率については、国立社会保障・人口問題研究所[14]によると、2023年には2.07であり、2000年以降この水準を保っている。また、1974年以降現在においても合計特殊出生率が人口置換率を下回っており、合計特殊出生率については1974年に2.05であったが1975年には1.91と2を切りそれ以降大きく回復することなく低下を続けている。

OECD[10]は、45～49歳において1人も子どもを産んでいない女性の割合は生涯無子率と定義している。OECD[10]では、1975年生まれの生涯無子率は28.3%で、これはOECD加盟国平均の16.2%を大幅に超えていることが示される。さらに、国立社会保障・人口問題研究所[14]によると、2005年生まれの女性が50歳になった時の無子割合は33.4%に達すると推計される。

以上より、国際的にみても急速な出生率の低下を経験しており、さらには子どもを持たない人の割合が非常に多い国であり、今後もその傾向には歯止めが利かないと予想される。このことから、少子化に関する議論は、子を持たない個人について「なぜ持たないのか／持てないのか」、子を持つ個人について追加で子を「なぜ持たないのか／持てないのか」を分けて考察する必要がある。本研究では、1 more baby 応援団 [16] がWEBアンケートにより実施した「夫婦の出産意識調査」の個票データを用いて、「子を持つか」および「追加で子を持つか否か」という出産の意思決定に影響を及ぼす要因について明らかにする。具体的には、無子個人と有子個人を区別し、それぞれについて出産意向に影響を与える要因を実証的に分析する。そのために、初産の意思決定と追加出産の意思決定の違いに注目し、両者に影響を与える要因を明らかにすることを通じて、少子化のメカニズムへの理解を深めることを目的とする。このように、本稿では、初産と2人目以降の出産の意思決定に注目するところに、本稿の特色がある。

2. 先行研究

出生行動に影響を及ぼす要因として、Becker [2]以降、経済学では出生率低下について様々な視点から研究が行われている。そこでは、個人の合理的な行動の結果、子を何人もつかという意思決定において、「子を持つ」ということが前提に分析が行われているものが非常に多い。一方、Momota[5]では、出生率が長期的に資本蓄積と社会厚生に及ぼす影響について、出生率を、1) 母

親が生む子どもの数(Intensive margin)、2) 母親率(Extensive margin)、とに区別して分析を行っている。一般に、経済学では出生率に関する議論は上記 1) が想定されることが多く、2) を導入することで初めての子を持つかどうか(初産)についての意思決定を組み込むことが可能となる。また、守泉[8]では、日本における無子割合の長期的な動向および女性を対象として無子の要因を、実証分析により明らかにしている。ここでは、無子を不妊や健康上の理由による無子である「非自発的無子」と結婚・出産の延期や無子を望む「自発的無子」に区別し、若い世代ほど自発的無子が増加していることを指摘している。また、無子志向型の無子女性について、結婚解消経験や子どものふれあい経験の少なさが、有意に影響を及ぼしていることも明らかにしている。

人口経済学の枠組みで指摘される出生行動に影響を及ぼす主要因として、まず所得があげられる。Becker[2]は、親が子どもを消費財とみなす場合、消費者の効用最大化行動の結果として最適な出生率が導出されることを示した。同モデルでは、所得が高いほど出生率も高くなる一方で、子育てに機会費用が発生する場合、母親の賃金率の上昇は出生率を引き下げる可能性が示される。このことから、子育てには時間がかかり、母親が中心となって育児を担っているのであれば、母親の賃金率上昇は出生率にネガティブな影響を及ぼし、一方で父親の所得増加は出生率に対してポジティブな影響を及ぼしうるということが分かる。

Doepke and Tertilt[3]では男性よりも女性のほうが子どもへの選好が強いことが示され、Attanasio and Lechene [1]では女性の所得と子どもの衣服や食料に対する支出額との間に正の相関があることが示されている。これらの研究は、女性を取り巻く社会的環境に出生率が依存しうることを示唆している。

宇南山[6]は、集合体モデルを用いて日本の未婚率上昇の要因について考察を行っている。同研究では、妻は結婚によってフルタイム労働者として就労できるかどうか外生的要因によって決定され、確率的にパートタイム労働者になるリスクが想定される。つまり、結婚という選択を行うことにより、妻の所得が低下するパートタイム労働者になるというコストを負担する状況が想定される。別府[7]では、有配偶女性について、出生率および出生前後の就業継続率が正規就業ではともに1987年から2007年にかけて上昇しているものの、非正規就業については両指標とも低水準で推移していることが示される。また、夫婦間の交渉において妻の交渉力が高くても、前述のリスクが存在する限り、女性が結婚するとは限らないことが示されている。これを出生率の枠組みに当てはめると、結婚において妻が

産後、子育て等の時間を含めた負担を負い、かつ労働市場への復帰が実現されないというリスクが存在するのであれば、子どもを産まないという選択肢は十分にありうることとなる。

以上を踏まえると、出生率に影響を及ぼす要因として、(1)経済要因(収入、経済的不安、物価、将来不安)、(2)就業・ワークライフバランス(本人の働き方、配偶者の働き方、労働時間・日数、育児休業制度・企業風土)、(3)家庭内要因・夫婦関係(バーゲニング、育児のシェア、家事分担への満足度、夫婦関係の良好さ、ワンオペ感)などが挙げられる。

このほか、出産後の女性の30-50%が経験するといわれる「マタニティブルー」も影響を及ぼしていると考えられる。厚生労働省[12]では、子育ては養育者に喜びや生きがいなどのポジティブな影響を及ぼす反面、子育て自体が心の健康にネガティブな影響を及ぼすストレス要因として作用する可能性が指摘される。ここから、(4)心理的・健康要因(出産・育児不安、健康・年齢不安、第1子出産時の経験)なども影響を及ぼす要因として作用すると考えられる。

また、Fluchtmann et al. [4]では、育児休業給付、保育制度、家族手当などの公的支出の拡大が出生率を引き上げることが示されている。このことから、(5)社会制度・地域環境(自治体の子育て環境)も重要な要因として位置づけられる。

3. データと方法

本研究の目的は、「子を持つか」および「追加で子を持つか否か」という出産の意思決定に影響を及ぼす要因について明らかにすることである。ただし、本稿では出産意欲と出産意向を特に区別せず、同じものとする。そのうえで、無子個人と既に子がいる有子個人に分けて、それぞれのグループの出産意向に影響を及ぼす要因を確認するために、以下では多変量解析を行う。

3.1 データ

研究では、1more Baby 応援団[16]の「夫婦の出産意識調査」の個票データを使用する。調査対象は既婚者2,961名で、子ありが1,974名、子なしが987名で構成

表1 子どもの数の割合

子どもの人数	割合 (%)
0	33.3
1	33.3
2	27.8
3	5.0
4	0.5
計	100

され、子の内訳は表 1 にて示される。

表 1 のとおり、33.3%が子どもを持たず、33.3%が子を 1 人、27.8%が子を 2 人、5.0%が子を 3 人、そして 0.5%が子を 4 人持つ。ただし、子どものいる既婚者の平均子ども数は 1.59 人である。

3.2 変数の定義

3.2.1 被説明変数：出産意向

本研究では、今後の出産意欲である出産の意向を説明変数とする。守泉[9]では、出生意欲（＝出産意欲）についての整理が行われている。そこでは、出生意欲の子どもの数に関するデータとして、「個人が目標とする子ども数」と「平均的な、あるいは典型的な家族に適した子ども数」に大別されること、また、出生意欲には子どもを持つ時期に関する意識も含むとしている。1 more baby 応援団[16]が実施した調査では、守泉[9]の出生意欲における後者に対応するデータがとられている。具体的には、「出産する／したいと思う（以降、出産意向あり）」「出産はしないと思う（以降、出産意向なし）」「わからない」と回答を求めているが、本研究では「わからない」を除いた「出産意向あり」を 1、「出産意向なし」を 0 として二値変数として再構成を行った。

3.2.2 主な説明変数：世帯年収

先行研究でも述べたように、出生行動は所得に依存するため、本研究でも世帯年収を主要な説明変数の一つとして用いる。ただし、調査票における世帯年収は 100 万円ごとの細かい所得階級で回答されている。本研究では、中所得者層を中心に、低所得者層や高所得者層の出産意向がどのように異なるかについても研究の主眼にしている。そのため、世帯年収を総務省統計局「家計調査」における収入五分位区分に基づき、次のように整理を行った。

まず、100 万円未満～400 万円未満を「低所得層」、400～500 万円未満を「中低所得層」、500～700 万円未満を「中所得層」、700～900 万円未満を「中高所得層」、900 万円以上を「高所得層」とした。そのうえで、中所得層を基準カテゴリとして、それぞれの所得階層についてダミー変数を作成して分析に投入している。

3.2.3 主な説明変数：年齢

年齢は連続変数であり、回答者の回答時点における実年齢（歳）であり、20 歳から 49 歳までの個人の回答があった。出産意向は年齢と非常に密接な関係があり、令和 6 年度においては厚生労働省[11]によると合計特殊出生率について、母の年齢（5 歳階級）別に見ると 30～34 歳で最も出生率が高く、25～29 歳がそれに続く。このことから本研究では、回答者の年齢に

ついて 1 歳刻みなどの細かな差を比較することよりも、年齢を 5 歳階級で見ることにより、年齢階級別の出生意向の差異を分析したい。具体的には、20～24 歳を「20 代前半層」、25～29 歳を「20 代後半層」、30～34 歳を「30 代前半層」、35～39 歳を「30 代後半層」、40～44 歳を「40 代前半層」、45～49 歳を「40 代後半層」とした。

3.2.4 主な説明変数：子どもダミー

出産意向については、4.1 にて後述するように、無子個人と有子個人とでは出産の意向およびそのタイミングが異なる傾向がみられる。さらに、子の数が個人の効用に影響を及ぼし、さらに子を消費財とみなす出生モデルにおいては、一般に効用関数は対数形が想定される。そのもとでは、個人は子の数に対して限界効用逓減が仮定される。つまり、子をたくさん持てば持つほど子から得られる効用水準は低下すると考えられる。そこで、本研究においては、後述の追加出産意向モデルにおいては、「子が 1 人いる」を基準カテゴリとして、「子が 2 人＝2」、「子が 3 人以上＝3+」として、モデルに投入している。

3.2.5 その他の説明変数

先行研究では、出生行動に影響を及ぼす要因として、①経済要因、②就業・ワークライフバランス、③家庭内要因・夫婦関係、④心理的・健康要因、⑤社会制度・地域環境といった観点から整理がなされてきた。本研究においても、この枠組みを踏まえ、各要因に関連すると考えられる変数を分析に投入している。

具体的には、経済要因として世帯年収を、就業・ワークライフバランスに関しては本人および配偶者の就業形態や職場環境に関する変数を用いている。また、家庭内要因・夫婦関係については、育児や家事の分担状況、夫婦関係の良好さに関する変数を含めている。さらに、子育てに対する不安や出産に対するためらいといった主観的評価は、心理的・健康要因として位置づけている。加えて、地域の子育て環境を示す変数や、年齢、結婚年数、子の数といったライフコース上の要因についても、統制変数として考慮している。

4. 分析結果

4.1. 記述統計

分析に用いる主な変数の平均／割合は表 2 の通りである。

表 2 分析に使用する主な変数の平均／割合

	平均	割合 (%)
出産の意向あり		54.3
出産の意向なし		31.8
出産の意向わからない		14.0
子ども数	1.1	
年齢 (歳)	32.9	
結婚年数 (年)	4～6年	
世帯年収	700～800万円未満	
低所得者層		21.9
中間層		67.3
高所得者層		10.8

※中間層は中低所得者層、中所得者層、中高所得者層の合計

表 3 は、出産の動向について「意志あり」と「意向なし」グループにおいて、実際の子ども数と出産のタイミングについて示したものである。ただし、各セル内は同じ子ども数を持つグループのうち、各々の出産の動向およびタイミングについて回答した個人の割合である。表 3 より、子どもを産んでいない無子個人について、出産の意向ありが 75.4%と出産意向を持つ割合が高く、その多くが「すぐに欲しい」の 56.0%であり、比較的早い段階での出産を希望している傾向がある。有子のうち「子が 1 人」の個人は、無子個人と同じく出産の意向を持つ割合は高いが、タイミングとしては「すぐ欲しい」と「1～2 年後」に分散されている傾向にある。「子が 2 人、3 人」の個人については、「出産の意向なし」が多数を占めている。また、「出産の意向あり」の個人についても、「すぐに欲しい」が最も多いものの、「1～2 年後」「3～5 年後」と無子および「子が

1 人」に比べると、出産のタイミングは後ろにずれており、出産時期を先送りする傾向がみられる。また、「いずれは欲しい」の回答が「子が 3 人」の個人には多く、子は欲しいものの現実味を帯びていない可能性がある。「子が 4 人」の個人は全体の 0.5%であるが、出産の意向は高くさらにそのタイミングも先送りする傾向がない事も確認できる。

4.2 初産および追加出産の意向の有無の規定要因について：二項ロジスティック回帰分析

前述の通り、本研究における被説明変数は出産意向について二値変数であり、二項ロジスティック回帰分析により、ある事象が生起する確率を世帯収入や第一子の有無などの説明変数によってモデル化する。出産意向については「初めての子を持つか否か」という意思決定と、すでに子がいて「次の子を持つか否か」と

表 3 「出産の意向あり・なし」グループの出産のタイミングと子ども数

子どもの数／タイミング	出産の意向あり (%)					出産の意向なし (%)	計
	すぐに欲しい	1～2年後	3～5年後	6年後以降	いずれは欲しい		
0人	56.0	12.7	2.8		3.9	24.6	100
1人	37.4	25.9	9.4		3.9	23.4	100
2人	14.7	8.7	6.8	0.4	6.5	62.9	100
3人	12.8	6.0	3.8		12.0	65.4	100
4人	46.7				6.7	46.7	100

表 4 初産意思決定および追加出産意思決定における二項ロジスティック回帰分析の結果

モデル		初産意思決定			追加出産意思決定		
		オッズ比	p	95% 信頼区間	オッズ比	p	95% 信頼区間
M ₀	(Intercept)	3.289	< .001	2.713 ~3.988	1.319	< .001	1.169 ~1.489
M ₁	(Intercept)	1.012	0.997	0.002 ~478.545	59.359	< .001	12.216 ~288.431
	年齢区分：30代前半(基準)						
	年齢区分：20代前半	4.264 × 10 ⁻⁶	0.988	0 ~ ∞	0.839	0.785	0.237 ~2.969
	年齢区分：20代後半	1.298	0.486	0.623 ~2.702	0.993	0.975	0.642 ~1.537
	年齢区分：30代後半	0.307	< .001	0.169 ~0.558	0.502	< .001	0.352 ~0.714
	年齢区分：40代前半	0.524	0.45	0.098 ~2.799	0.305	0.004	0.137 ~0.679
	年齢区分：40代後半	0.197	0.155	0.021 ~1.843	0.204	0.027	0.049 ~0.838
	結婚年数	0.485	< .001	0.358 ~0.659	0.509	< .001	0.416 ~0.622
	世帯年収：中所得(基準)						
	世帯年収:低所得	1.175	0.691	0.53 ~2.605	1.144	0.571	0.718 ~1.823
	世帯年収:中低所得	0.723	0.444	0.314 ~1.661	0.837	0.465	0.521 ~1.347
	世帯年収:中高所得	1.066	0.86	0.524 ~2.168	0.97	0.888	0.637 ~1.477
	世帯年収:高所得	0.823	0.672	0.333 ~2.031	0.853	0.547	0.509 ~1.431
	家庭幸福	1.003	0.651	0.991 ~1.015	1.019	< .001	1.01 ~1.028
	育児共有				1.202	0.334	0.828 ~1.745
	夫婦関係	0.424	0.026	0.199 ~0.903	0.461	< .001	0.31 ~0.685
	家事分担	0.701	0.187	0.414 ~1.188	1.097	0.599	0.776 ~1.552
	上司理解	0.79	0.385	0.465 ~1.344	1.012	0.946	0.72 ~1.422
	地域育児	0.782	0.196	0.539 ~1.135	0.759	0.016	0.605 ~0.95
	1人目出産ためらい	2.483	< .001	1.925 ~3.203	0.943	0.469	0.805 ~1.105
	子育て不安	0.75	0.132	0.516 ~1.091	0.832	0.111	0.664 ~1.043
	性別：男(基準)						
	性別：女	0.487	0.097	0.208 ~1.14	0.95	0.818	0.616 ~1.466
	本人就業 正規職員・従業員(基準)						
	本人就業(経営者・役員)	3.240 × 10 ⁺⁷	0.992	0 ~ ∞	1.03	0.961	0.322 ~3.287
	本人就業(契約社員・派遣社員・嘱託社員)	0.466	0.092	0.192 ~1.131	0.691	0.307	0.34 ~1.404
	本人就業(自営業)	0.684	0.801	0.036 ~13.074	1.097	0.893	0.284 ~4.245
	本人就業(個人事業主)	1.272	0.772	0.25 ~6.485	3.252	0.078	0.875 ~12.088
	本人就業(パート・アルバイト)	0.569	0.094	0.294 ~1.101	0.686	0.067	0.458 ~1.027
	配偶者就業(正規職員・従業員)(基準)						
	配偶者就業(経営者・役員)	0.191	0.022	0.046 ~0.786	1.045	0.921	0.439 ~2.487
	配偶者就業(契約社員・派遣社員・嘱託社員)	0.608	0.402	0.19 ~1.946	0.597	0.246	0.25 ~1.427
	配偶者就業(自営業)	0.502	0.475	0.076 ~3.33	1.258	0.65	0.467 ~3.391
	配偶者就業(個人事業主)	0.524	0.38	0.124 ~2.217	0.284	0.061	0.076 ~1.062
	配偶者就業(パート・アルバイト)	0.602	0.344	0.211 ~1.721	0.84	0.551	0.474 ~1.49
	テレワーク_本人(積極的)(基準)						
	テレワーク_本人(時々)	0.793	0.799	0.133 ~4.714	1.087	0.86	0.428 ~2.76
	テレワーク_本人(していない)	0.703	0.646	0.156 ~3.163	0.97	0.94	0.434 ~2.166
	テレワーク_本人(今はしていない)	0.677	0.65	0.125 ~3.657	1.128	0.807	0.428 ~2.972
	テレワーク_配偶者(積極的)(基準)						
	テレワーク_配偶者(時々)	3.095	0.242	0.466 ~20.563	0.624	0.33	0.241 ~1.614
	テレワーク_配偶者(していない)	2.158	0.329	0.461 ~10.114	0.513	0.12	0.221 ~1.19
	テレワーク_配偶者(今はしていない)	1.75	0.533	0.302 ~10.154	0.734	0.537	0.275 ~1.959
	子の数：1人(基準)						
	子の数:2人				0.21	< .001	0.151 ~0.293
	子の数：3人以上				0.209	< .001	0.123 ~0.354

いう意思決定、という異なる意思決定の段階があり、本節ではこれらの意思決定の段階の違いに注目し、それぞれにおいて出産の意思決定にどのような要因が影響を及ぼしているかについて検討したい。

本節において用いるロジスティック回帰分析の基本的な推計式は、以下の通りである。

$$\log\left(\frac{p}{1-p}\right) = \text{logit } p = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots$$

ここで、 p は出産の意向ありの生起確率であり、説明変数は 3. 2 節で示した通りである。

4.2.1 初産における出産の意思決定について：初産意向モデル

本項では、無子個人を対象にし、初めての出産に対する意向に影響を及ぼす要因について確認をした。その分析の結果は表 4 にてまとめられるが、初産の意向モデルにおいては、

$\chi^2(544) = 215.741, \quad p < .001, \quad \text{McFadden}R^2 = .343$
から十分なモデルの当てはまりが確認された。

初産の意向について、年齢については 30 代後半群については有意に低下しており（OR=0.159）、40 代後半群についてはさらに低下している（OR=0.03）。結婚年数についても、結婚年数が長いほど、初産の意向は有意に低下し（OR=0.469）、1 人目の出産にためらいを感じている人ほど有意に初産の意向は低下している（OR=2.507）ことが分かった。

4.2.2 2 人目以降の出産における意思決定について：追加出産意思決定モデル

本項では、有子個人を対象にして、追加で子（第 2 子以降）を産むか否かの意思決定に注目する。基本的なモデルは前項と同様であるが、本節では追加の説明変数として育児共有「子育ての大変さをしっかりと夫婦でシェアできている」を投入する。さらに、すでに子がいて追加でもう一人出産するか否か、という意思決定において、子が 1 人いる状態から「もう一人産むか否か」、と子が 3 人いる状態から「もう一人産むか否か」、とでは追加出産としての意味が違わずである。

そこで、本節では子の数について「子が一人いる」を基準カテゴリとして、「子を 2 人いる＝2」、「子が 3 人以上いる＝3+」としたうえで、カテゴリ化してモデルに投入している。分析の結果は表 4 にてまとめられる。

分析の結果、

$\chi^2(1031) = 370.795, \quad p < .001, \quad \text{McFadden}R^2 = .254$
から十分なモデルの当てはまりが確認された。

また、年齢区分については、基準の 30 代前半群に比べると 30 代後半群は有意に追加出産の意向が低下

表 5 初産意思決定と追加出産決定の比較表

	初産	追加出産
年齢	●	●
経済要因	－	－
就業・ワークライフバランス	－	－
家庭内要因・夫婦関係	－	●
心理的・健康要因	●	－
社会制度・地域環境	－	●

しており（OR=0.502）、この傾向は 40 代前半群、40 代後半群と年齢を重ねるほど追加出産の意向が追加しているがわかる。また、結婚年数が長いほど、夫婦関係が円満であるほど、有意に追加出産の意向は低下している（それぞれ、OR=0.509、OR=0.461）。地域育児については、住んでいる地域において子育てがしにくいと感じるほど、有意に追加の出産意向は低下している（OR=0.759）一方で、家庭が幸福であるほど追加出産の意向は高まることが確認できる（OR=1.019）。

子の数については、子が 1 人いる個人に比べると、子が 2 人いる個人および 3 人以上いる個人の追加出産の意向はいずれも有意に低下していることも分かった（それぞれ、OR=0.21、OR=0.209）。

4.3 初産意思決定と追加出産意思決定の比較

ここまで、子どもを持たない個人を対象にした初産意思決定と、子どもを持つ個人を対象とした追加出産意思決定について、出産の意向について「あり＝1」「なし＝0」の二値変数を用いて二項ロジスティック回帰分析を行った。その結果をまとめたものが以下の表 5 であり、有意であったものに「●」、非有意に「－」をつけている。

表 5 からも見て取れるように、初産の意思決定および追加出産の意思決定において、年齢は非常に重要な要素であり、年齢を重ねるほど出産の意向は低下していた。ただし、追加出産の意思決定においては、40 歳代を通じてその傾向が有意に確認されていた。一人目出産である初産の意思決定においては、1 人目出産のためらいが出産の意向を押し下げる要因であったのに対して、追加出産においては家庭内要因・夫婦関係、社会制度・地域環境およびすでに何人子を持っているか、という要因が影響を与えていることが分かった。

5. 考察と結論

本研究では、2023 年の夫婦の出産意識調査データを用いて、出生行動に影響を及ぼす要因を実証的に検討した。分析の結果、以下が明らかになった。

第一に、初産の意思決定においては、年齢や経験したことの無い出産に対する不安が大きく影響を及ぼすことが分かった。第二に、追加出産については年齢に加えて、子どもを育てるための環境に関連する要因に影響を受けることが分かった。これらの結果は、初産と追加出産とでは、意思決定において異なる制約や価値判断が介在していることが示唆され、追加出産については、実際の子育てにおいて発生する負担などが影響していると考えられる。

一般に、経済学における出生理論では、所得水準などの経済的制約が出生行動に影響を及ぼす要因としてあげられる。本研究では、少なくとも初産意向においては、年齢等を一定とした多変量分析において、所得階層間で明確な差は確認されなかった。この結果は、初産の意思決定において、子育てにかかる費用が具体的に認識される以前の段階で、意思決定を行っている可能性を示唆している。追加出産においても、所得階層間で明確な差は確認されなかったが、実際の子育て経験を通じて子育て費用が可視化されており、所得水準そのものよりも、子育てに関する負担が出産意向とより密接に関連している可能性がある。

本研究の結果から、所得階層間の世帯年収の差異のみでは、出産の意向の違いを十分に説明できない可能性が示唆された。このことは、所得水準の引き上げのみを目的とした政策は、必ずしも出生率の回復に直結しない可能性があることを示している。さらには、初産と追加出産では、少子化対策としての出産支援政策を一律で設計するのではなく、出生段階に応じて異なるアプローチが必要であることを示している。

こども家庭庁[15]では、2030年代に入ると日本の若年人口は現在の倍速で急減するため、「日本のラストチャンス」として集中的に実施する具体的な政策を「こども・子育て支援加速化プラン」とし、ライフステージ別（結婚、妊娠～出産・産後、就学前、小・中学校、高校、大学等）のプランとして、次のような政策が示されている。

経済的支援の拡充として、児童手当の所得制限撤廃と支給期間の高校生年代までの延長、高等教育の無償化拡大などが実施されている。保育サービスの充実では、「こども誰でも通園制度」の創設により未就園児の定期的な預かりサービスや保育士処遇の改善などが進められている。働き方改革・両立支援としては、男性育休の取得促進、育児休業給付の引き上げ（最大手取り割相当）などがあり、妊娠・出産支援では、不妊治療の保険適用拡大、伴走型相談支援の充実、産後ケア事業などの促進が推進されている。これらの中でも、特に初めての出産を後押しするような環境整備や出産そのものの不安を軽減する政策により、一人目に対す

る政策として有効であり、実際の子育ての負担を軽減するような支援や、家庭内・社会的なサポート体制の充実が、より重要な役割を果たす可能性があり、これらの異なる効果を意識した少子化対策を周知することも有効であると思われる。

研究にはいくつかの残された課題がある。本研究ではクロスセクションデータを用いているため、因果関係を厳密に特定することは難しい。所得や家族関係が出産意向に影響を及ぼしているのか、あるいは出産意向の違いがこれらの要因に影響を及ぼしているのかについては、明らかにすることは有用であると考えられる。

謝辞

本分析を行うにあたり、東京大学社会科学研究所附属社会調査データアーカイブ研究センターSSJ データアーカイブから「夫婦の出産意識調査, 2023」(1 more Baby 応援団) の個票データの提供を頂きました (<https://doi.org/10.34500/SSJDA.1624>)。ここに記して謝意を表します。

参考文献

- [1] Attanasio, O., and V. Lechene (2002) “Tests of Income Pooling in Household Decisions,” *Review of Economic Dynamics*, vol.5, issue 4, pp.720-748.
- [2] Becker, G.S. (1960) “An Economic Analysis of Fertility,” *Demographic and Economic Change in Developed Countries*. Columbia University Press.
- [3] Doepke, M., and M. Tertilt (2009) “Women's Liberation: What's in It for Men?,” *The Quarterly Journal of Economics*, vol.124, issue 4, pp.1541-1591.
- [4] Fluchtmann, J., van Veen, T., and W. Adema (2023) “Fertility, employment and family policy: A cross-country perspective,” *OECD Papers on Social, Employment and Migration*, No. 286, OECD Publishing, Paris.
- [5] Momota, A. (2016) “Intensive and Extensive Margins of Fertility, Capital Accumulation, and Economic Welfare,” *Journal of Public Economics*, 133, pp.90-110, January 2016.
- [6] 宇南山卓(2014)「女性の労働市場・家計内分配と未婚化」RIETI Discussion Paper Series 14-J-048.
- [7] 別府志海(2012)「有配偶女性の就業移動と出生力」人口問題研究 第68巻第1号 1-13頁.
- [8] 守泉理恵(2019)「日本における無子に関する研究」人口問題研究 第75巻第1号 26-54頁.
- [9] 守泉理恵(2025)『日本における出生意欲と出生数の乖離からみた少子化の要因：産み育てやすい社会実現のための政策的対応に関する研究』博士論文、お茶の水女子大学.
- [10] OECD(2024) Society at a Glance 2024 - Country Notes: 日本
<https://www.oecd.org/ja/publications/2024/06/society-at-a-glance-2024-country->

notes_d98f4d80/japan_e8ace1fa.html (2026 年 1 月 25 日閲覧)

- [11]厚生労働省(2024a)「人口動態統計月報年計(概数)の概況」

<https://www.mhlw.go.jp/toukei/saikin/hw/jinkou/geppo/nengai24/dl/gaikyouR6.pdf> (2026 年 1 月 16 日閲覧)

- [12]厚生労働省(2024b)『厚生労働白書 令和 6 年版』

- [13]国立社会保障・人口問題研究所(2021)「第 16 回出生動向基本調査(夫婦調査)」

https://www.ipss.go.jp/ps-doukou/j/doukou16/JNFS1-6_ReportALL.pdf
(2026 年 1 月 16 日閲覧)

- [14]国立社会保障・人口問題研究所(2023)「日本の将来推計人口(令和 5 年推計)」

https://www.ipss.go.jp/pp-zenkoku/j/zenkoku2023/pp2023_ReportALLc.pdf
(2026 年 1 月 16 日閲覧)

- [15]こども家庭庁(2023)『令和 6 版こども白書』

<https://www.cfa.go.jp/resources/white-paper/r06> (2026 年 1 月 25 日閲覧)

- [16] 1 more Baby 応援団(2023)『夫婦の出産意識調査』SSJ データアーカイブ、東京大学社会科学研究所附属社会調査データアーカイブ研究センター。

<https://ssjda.iss.u-tokyo.ac.jp/Direct/gaiyo.php?lang=jpn&eid=1624>
(2026 年 2 月 11 日閲覧)

就職活動における情報探索行動と行動経済学的個人特性の関係分析

後藤 侑美[†] 山本 修平^{††}

[†] 筑波大学情報学群知識情報・図書館学類 〒305-8550 茨城県つくば市春日 1-2

^{††} 筑波大学図書館情報メディア系 〒305-8550 茨城県つくば市春日 1-2

E-mail: [†]ts2211865@u.tsukuba.ac.jp, ^{††}syamamoto@slis.tsukuba.ac.jp

あらまし 日本では新規学卒者の就職活動の早期化と長期化が顕著である。この現状は、政府の定める採用活動日程ルールから大きく乖離しており、学業との両立を鑑みると、就職活動の効率化支援が喫緊の課題となっている。これまでの就職活動研究では、「意思決定理論」や「進路探索行動（CES）」の文脈で、情報探索行動が分析されてきた。特に情報探索行動の内容と結果の関連が明らかにされてきたが、個人差の要因については言及されていなかった。本研究では、この個人差を明らかにするため、就職活動を労力と将来的な利益を比較する異時点間選択の意思決定と捉え、行動経済学的な個人特性である時間選好とリスク選好に着目する。

キーワード 就職活動、情報探索行動、時間選好、質問紙調査

1 はじめに

就職活動の現状：昨今、日本では新規学卒者の就職活動の早期化が問題となっている。内閣府が実施した調査によれば、企業説明会等への参加時期のピークが卒業年次前年の9月以前と答えた学生の割合は2024年度卒業・修了予定者では27.5%であり、2020年度卒業・修了予定者の約10%と比較すると大幅に増加していることがわかる[18]。また、就職活動に要する期間についても、「9か月程度以上」と回答する学生の割合が2020年度卒業・修了予定者では30.0%であったのに対し、2024年度卒業・終了予定者では48.3%となっており、長期化も顕著になっていることが明らかである[18]。これらの現状は、政府が経済団体などに対して要請している「卒業年次前年の3月1日以前に広報活動を開始しないこと」や「卒業年次前年の6月1日以前に採用選考活動を開始しないこと」といった採用活動日程ルールから大きく乖離するものである[17]。こうした採用活動日程ルールの形骸化が進む中、制度的な規制によって早期化を是正することは現実的ではなく、むしろ早期化している就職活動に直面する学生一人ひとりに対する実践的な支援が求められる。特に学業等との両立も鑑みると、就職活動の効率化を支援していくことが喫緊の課題である。

就職活動研究の概観：ここで重要となるのは、就職活動における具体的な行動の傾向を分析することである。これまでの日本の就職活動に関する研究は様々なアプローチで行われてきたが、中でも具体的に表出される行動を扱ってきたのは「意思決定理論」や「進路探索行動（Career Exploration Survey; CES）」の文脈である[14]。前者は、就職活動を意思決定のプロセスとそれに伴う情報処理の過程として捉える理論であり、後者は職業的な意思決定をする中で、個人が職業や自己に関する情報を収集することを指す[7]。この2つの文脈の中でも、特に着目されているものが情報探索行動である。特にCESは自己に関する探索（self-exploration）と環境に関する探索（environmental

exploration）の2側面から構成されており、就職活動における情報探索行動を捉える尺度として多く用いられてきた。具体的な研究の例として、下村らは大学生約50名を対象に質問紙を用いた縦断調査を行い、就職活動においてOB/OGや就職サイトといった情報源がそれぞれどのような情報内容を提供しているか、またそれぞれの情報源の利用実態が就職活動の結果にどのような影響を与えているかについて分析した[11]。また、矢崎らは前述の進路探索行動尺度が12項目のみと、やや抽象的であることを指摘し、下村らの情報源に加えて「何を調べたか」、すなわち情報探索内容も併せて調査する「就職に関する情報探索行動尺度」を作成した[19]。

先行研究の課題：しかし、これらの先行研究は情報探索行動の内容と就職活動の結果との関連を明らかにすることにとどまっており、情報探索行動やその結果との相関の個人差については述べられていない。本研究の目的である個々人に対する支援を行っていくうえで、その内容を個人に最適化することは重要であり、そのためには就職活動における情報探索行動の個人差を分析することは欠かせない。ここで、就職活動に限らず、広く情報探索行動を扱う先行研究を調査すると、Big-Fiveなどに代表される心理学的な個人特性が情報探索行動に有意に影響を与えていることがわかる[9]。このことを鑑みると、就職活動における研究においても、同様に個人特性の影響を調査することは有用であると考えられる。また、この点が明らかになることで、学生に一律の支援を行うのではなく、個人ごとに最適化された支援を行うことも可能になるであろう。

行動経済学的アプローチの有用性：そこで、就職活動における学生の行動を予測する個人特性として、行動経済学において用いられる時間選好とリスク選好に着目する。これまでの就職活動の研究において、関連する変数としてよく用いられてきたものには、社会認知理論の中心概念である「自己効力」などがある[14]。しかし、本研究で用いる2つの行動経済学的個人特性は、これらの先行研究と比較して、より潜在的な変数であり、かつ、就職活動における行動を広範に捉えることができると考

えられる。

この理由として、就職活動を「異時点間選択」として捉えられること、また就職活動は不確実性を多分に孕んだ意思決定であることが挙げられる。就職活動を意思決定として捉え、その特徴を考えると、就職活動は労力や金銭・時間的負担といった現在のコストと、キャリア形成や給与のような将来的な利益を天秤にかけると、異時点間選択として捉えることができる。また、就職活動によって得られる将来の利益は先述のキャリアのように、不確実性を多分に孕むものであるともいえる。このような意思決定としての特徴を踏まえると、将来の利益を割り引く程度である時間選好や、個々人のリスクに対する態度を測定するリスク選好は就職活動における行動をより良く説明すると思われる [5] [4]。

時間選好やリスク選好が就職活動に影響を与えると考えられる具体的なケースを挙げる。時間選好が高い (現在志向的な) 学生は将来の利益を低く見積もり、目先の労力を回避する傾向があるため、情報探索が不足し、入社企業とのミスマッチが起こりやすくなると想定される。同様に、リスク愛好的な学生は少数の志望度の高い企業に絞り込み、限定的な情報探索で選考に臨む可能性がある。このように、意思決定としての特徴を考慮した個人特性を用いることで、その個人特性が情報探索行動や結果を予測することが期待できる。

本研究の目的: このような背景から、本研究では就職活動における情報探索行動の個人差を、行動経済学的な個人特性である時間選好とリスク選好を用いて分析する。さらに、これらと就職活動の結果を併せて分析することで、学生が自らの特性に応じて効果的に就職活動を進めるための支援の示唆を得ることを目指す。主なリサーチクエスチョン (RQ) は以下の通りである。

- RQ1: 行動経済学的な個人特性によって、就職活動において利用する情報源は異なるか?
- RQ2: 行動経済学的な個人特性によって、就職活動において探索する情報内容は異なるか?
- RQ3: 行動経済学的な個人特性は、就職活動の効率化を学生個々人に対して支援する際の有用な手掛かりとなるか?

2 関連研究

本章では行動経済学的個人特性と実際の日常行動との関連を調査した論文として、盛本による「時間割引率および現在バイアス性がオンラインゲーム内コンテンツへの課金行動に与える影響」 [15] を示す。この研究では、オンラインゲームにおける「課金」行動に焦点を当てている。盛本はこの課金行動を「消費行動」として捉え、行動経済学的個人特性である時間割引率・現在バイアス・ナীবを用いて分析した。ここでは特に、本研究と強く関連する時間割引率と現在バイアスについて取り上げる。

時間割引率と現在バイアスは 1 章で述べた時間選好の指標であり、以下に示すような選択課題を用いて測定される [3]。

Q1: あなたは、ある金額をもらえることになりました。今日か 7 日

後にもらえますが、金額が異なります。もらえる日にちと金額について以下の選択肢「A」または「B」があれば、どちらを選びますか。1 から 9 の全ての組み合わせについて「A」または「B」の二択で回答してください。

	選択肢 A	選択肢 B
1	今日 3,005 円もらう	7 日後 3,014 円もらう
2	今日 3,003 円もらう	7 日後 3,297 円もらう
3	今日 3,008 円もらう	7 日後 3,037 円もらう
4	今日 3,000 円もらう	7 日後 3,000 円もらう
5	今日 3,005 円もらう	7 日後 5,951 円もらう
6	今日 3,009 円もらう	7 日後 3,068 円もらう
7	今日 3,001 円もらう	7 日後 3,119 円もらう
8	今日 3,002 円もらう	7 日後 2,996 円もらう
9	今日 3,008 円もらう	7 日後 3,011 円もらう

Q2: あなたは、ある金額をもらえることになりました。90 日後か 97 日後にもらえますが、金額が異なります。もらえる日にちと金額について以下の選択肢「A」または「B」があれば、どちらを選びますか。10 から 18 の全ての組み合わせについて「A」または「B」の二択で回答してください。

	選択肢 A	選択肢 B
1	90 日後 3,000 円もらう	97 日後 3,118 円もらう
2	90 日後 3,006 円もらう	97 日後 3,000 円もらう
3	90 日後 3,000 円もらう	97 日後 3,009 円もらう
4	90 日後 3,007 円もらう	97 日後 3,301 円もらう
5	90 日後 3,006 円もらう	97 日後 3,035 円もらう
6	90 日後 3,002 円もらう	97 日後 3,005 円もらう
7	90 日後 3,001 円もらう	97 日後 5,955 円もらう
8	90 日後 3,007 円もらう	97 日後 3,001 円もらう
9	90 日後 3,007 円もらう	97 日後 3,066 円もらう

※ Q1, Q2 ともに、実際の回答において、組み合わせの順番はランダムに並び替えられている。

Q1 はどの程度の金利がつけば 7 日間待つことができるかを測定し、「せっかちさ」の指標である時間割引率を導く。また、Q2 は 1 年後の時間割引率を測定する設問であり、この未来の時間割引率が Q1 の現在の時間割引率と一致しないことを「時間非整合」と呼ぶ。この時間非整合性の程度を「現在バイアス」として表す。現在バイアスが強く、先延ばし傾向のある人は Q2 のような「計画時点」での時間割引率が小さく、Q1 のような「実行時点」での時間割引率が多い。この場合、計画時点では我慢強いにもかかわらず、実際にその時点が訪れるとせっかちになり、目の前の小さな 7 日後の金利ではなく、目の前の小さな報酬を選択してしまう。

盛本はオンラインゲームにおける課金を「早くシナリオを進めたい」「早く育成したい」といった衝動性の伴う行動であると解釈し、時間割引率との関連が見られるとの仮説を立てた。また、現在バイアスに関してもオンラインゲームの「無料で始められる」「後天的に課金が発生する」といった特徴と照らし合わせ、課金と時間非整合性に関連が見られると考えた。加えて、「ガチャ」と呼ばれるギャンブル性を備えたオンラインゲームのシステムにも着目し、危険回避度 (リスク選好と同意) と課金行動の関連も検証した。リスク選好も同様に次のような選択

課題から求められる [4].

Q3：50%の確率で 10 万円が当たる「スピードくじ」があります。当たれば、賞金は今日すぐに支払われます。外れた場合、賞金はゼロです。あなたは「スピードくじ」をいくらなら買いますか。次の各値段について、それぞれ買う場合は「1」を、買わない場合は「2」の二択で回答してください。

[10 円,2,000 円,4,000 円,8,000 円,15,000 円,25,000 円,35,000 円,50,000 円]

盛本はこれらの仮説を検証するために、時間割引率・現在バイアス性・リスク選好に加えて、オンラインゲームの経験・オンラインゲームにおける課金行動の有無・所得や生活習慣に関する項目を含んだ質問紙調査を行った。結果として現在バイアスの強さと課金行動の関係には有意差が見られ、先延ばし傾向が強いほど課金しやすいことがわかった。このほかにも、時間非整合性への自覚の有無を表すナীব性¹と課金行動や喫煙行動と課金行動の間に関係が見られた。さらに、盛本は人々の行動をより良い方向へ導くための介入である「ナッジ」の提案を行っている [8]。具体的には、過剰な課金を防ぐために「課金上限金額の設定」や「上限金額の変更を複雑なプロセスにすること」が提案されている [15]。

3 手 法

3.1 調査の概要

就職活動における情報探索行動と行動経済学的な個人特性との関係を調べるため、本研究では 2025 年 9 月から 2025 年 10 月にかけて、学生を対象とした質問紙調査を行った。調査はオンライン上で行われ、リクルーティング・調査の実施にはセルフ型アンケートツール Freeasy を使用した。また、本調査は筑波大学図書館情報メディア系によって行われた倫理審査の承認を得ている (通知番号第 25-48)。この調査では、まず始めに 21 歳-26 歳の学生 1629 名に向けてスクリーニング調査を実施した。このうち「調査時点で大学 4 年生または修士課程 (博士前期課程) 2 年生である」「2026 年度に新規学卒者として入社する」「1 年以内に就職活動をした経験があり現在は行っていない」と回答し、かつ時間選好のアンケートに不備なく答えている者のみを対象に本調査を実施し、139 名の回答を得た。

3.2 質問紙の構成

3.2.1 スクリーニング調査

スクリーニング調査の内容を表 1 に示す。時間選好のアンケートは金利がマイナスの選択肢を選択した回答者、時間割引が双曲割引にも指数割引にも当てはまらない回答者を分析対象者から除いた。分析対象となった実験参加者については、Q1 において選択肢を金利順に並び替えた際に A から B へ変わった番号を時間割引率とし、Q1 と Q2 における切り替わりの番号の比を現在バイアスとした。なお、指数割引とは Q1 と Q2 の金利が同等である、すなわち、近い将来も遠い将来も同等にお金を割り引いていることを指し、双曲割引とは Q1 の年利よりも Q2 の年利が低い、すなわち、近い将来より遠い将来を小さく割り引くことを指す [1]。

表 1 スクリーニング調査の内容

設問内容 (回答方法)	
1	性別 (女性, 男性)
2	学年 (4 年, 修士 2 年, どちらにも当てはまらない)
3	2026 年に大学または大学院を卒業し、企業に就職するか (はい, いいえ)
4	就職活動の経験 (1 年以内に就職活動をした経験があり現在は行っていない, 現在就職活動を行っている最中である, 就職活動を行った経験がない)
5	時間選好 (2 章を参照)

表 2 項目の変更 (情報源)

追加	SNS, Web 上の口コミ, 生成 AI, インターンシップ, 就職ガイダンス (オンライン)
削除	就職情報雑誌, 企業から送付された資料

表 3 項目の変更 (情報探索内容)

追加	各企業の職種とその業務内容, 各企業でのワークライフバランス, 自分の幼少期や過去
削除	仕事の厳しさ, 企業側の電話の対応の仕方, スケジュール管理の仕方, 電話のかけ方企業側の説明の仕方, 面接の受け方, 一般常識, 敬語の使い方, 手紙の書き方, 服装

3.2.2 本 調 査

本調査の質問紙は「就職活動における情報探索行動」「行動経済学的個人特性」「就職活動の状況」の 3 領域で構成され、これらの関係性を明らかにすることを目的として作成された。

就職に関する情報探索行動：この部分では、下村らの研究 [11] と、それに基づいて矢崎らが作成した「就職に関する情報探索行動尺度」を参照した [19]。この尺度は 7 つの情報源と 44 の情報探索内容を組み合わせた 91 項目で構成されており、各項目の回答方法には 5 件法を用いている (1：全く行わなかった 6：とても行った)。しかし、2007 年に作成されている点や、情報源と情報探索内容の組み合わせが調査者によって独自に定められている点から、就職活動における情報探索行動の実態をより正確に測るための改善が必要であると考え、以下のような変更を加えた。第一に、2025 年現在の就職活動に沿った質問項目を設定するために、学生に対する調査を参考に情報源と情報探索内容の追加・削除を行った [12] [13]。追加・削除した項目を表 3 に、質問紙調査で利用した項目一覧を表 4、表 5 に記載する。

また、各情報探索内容においてどのような情報源が利用されているかについて調査するため、設問方法を次のように変更した。まず、12 項目の情報源と 36 項目の情報探索内容の利用度をそれぞれ独立して尋ねた。選択肢には 6 件法を用いた (1：全く行わなかった 6：頻繁に行った)。次に 36 項目の情報探索内容それぞれについてどのような情報源を利用したかを尋ねた。情報探索内容ごとに利用した情報源は複数回答できるようにし、「当てはまるものがない」「その他」の回答も設けた。なお、論文中の表で一部略称ラベルを使用することがあるため、その対応についても記載する。

行動経済学的個人特性：本研究では行動経済学的個人特性のうち、時間選好とリスク選好を用いた。時間選好に関してはス

表 4 情 報 源

略称ラベル	項目名
hp	企業の HP にアクセスした
site	就職サイト（リクナビ・マイナビなど）を使用・閲覧した
guidance_online	オンラインの就職ガイダンスに参加した
guidance_real	対面の就職ガイダンスに参加した
friend	友人と就職活動について話した
self	インターネットを用いたり、他者とのコミュニケーションを行わず、自分自身で思考を巡らせた
OBOG	OB/OG 訪問をした
senior	大学の先輩に就職活動に関する質問をした
sns	SNS で就職活動に関する情報を閲覧・検索した
review	WEB 上で就職活動に関する口コミ情報を閲覧・検索した
ai	生成 AI と就職活動について話した
internship	インターンシップに参加した

表 5 情報探索内容

略称ラベル	項目名
bc	各企業の事業内容を調べた
occupation	各企業の職種とその業務内容を調べた
future	各企業の方針・今後の動向を調べた
salary	各企業の給与を調べた
culture	各企業の社風を調べた
transfer	各企業の転勤の有無を調べた
career	各企業でのキャリアアップについて調べた
benefit	各企業の福利厚生について調べた
large	各企業の規模・知名度について調べた
labor_condition	各企業の労働条件について調べた
WLB	各企業でのワークライフバランスについて調べた
place	各企業の勤務地について調べた
ricruit_plan	各企業の採用計画について調べた
HR	各企業の人事の人の印象について調べた
atmosphere	各企業の職場の雰囲気について聞いた
OBOG	OB/OG の訪問の仕方について調べた
ES	ES の書き方について調べた
news	時事問題を調べた
manner	就職活動をする上でのマナー・礼儀について調べた
impression	実際に働いている人の印象を聞いた
search_company	企業研究の仕方を調べた
selfPR	自己 PR の仕方を調べた
society	社会・政治の動向を調べた
economics	経済の動向を調べた
hiring_trends	求人の動向を調べた
international	国際関係の動向を調べた
industry	業界の動向を調べた
capability	自分の能力について考えた
personality	自分の性格について考えた
aptitude	自分の適性について考えた
childhood	自分の幼少期や過去を振り返った
family	家庭環境を振り返った
interest	自分の興味・関心について考えた
want	自分のやりたいことについて考えた
major	自分の専攻と職種との関連について考えた
paren	親の希望を考えた

クリーニング調査の部分を参照されたい。リスク選好の測定に用いた設問は 2 章で示したものと同様である。なお、時間割引

表 6 就職活動に関する項目

内容	回答方法
就職活動に対する満足度	6 件法
入社予定企業への満足度	6 件法
入社予定企業が第 1 志望か	はい/いいえ
進路状況	就職先が決まっている/就職先を選びきれていない
就職活動を始めた時期	年月
就職活動を終了した時期	年月
1 つ目の内定を獲得した時期	年月

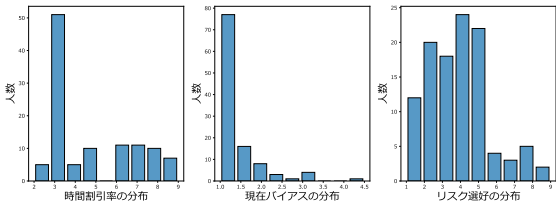


図 1 各選好の人数の分布。左から、時間割引率、現在バイアス、リスク選好の人数分布である。

率と同様に回答が「買う」から「買わない」へ切り替わった項目番号を個人のリスク選好の値とし、一度回答が「買わない」に切り替わった後に、再び「買う」に戻っている回答者は分析対象外とした。

就職活動に関する項目：個人の就職活動の状況・結果について尋ねた。質問内容は右の表のとおりである。

4 結 果

4.1 データ収集結果

4.1.1 分析対象者に関する情報

本調査回答者 139 名のうち、アンケートに不備なく回答している回答者のみを抽出し、110 名の回答を分析対象とした。分析対象者の性別の内訳は、男性が 25 名、女性が 85 名であった。また、学年の内訳は 4 年が 97 名、修士 2 年が 13 名であった。

また、分析対象者の時間割引率・現在バイアス・リスク選好の分布は以下のグラフ 1 の通りである。時間割引率は値が大きいほど、7 日後の金利を含んだ大きな報酬よりも目の前の小さな報酬を選択しやすいため、時間割引率が高いほどせっかちであると言える。また現在バイアスは現在時点の時間割引率と 90 日後の時点での時間割引率の比（現在時点の時間割引率/90 日後の時間割引率）であるため、値が大きくなるほど現在バイアスが強く、時間非整合的な選択をしやすいため、先延ばし傾向の強い学生であると言える。また、リスク選好は値が大きくなるほどリスク愛好的であることを示す。

4.1.2 尺度の検討

本研究では、矢崎らの「情報探索行動尺度」[19]をもとに、一部の項目を変更した尺度を用いた。ここで、この変更を加えた尺度の妥当性を検証するために、各項目の平均利用度を調べた。情報源・情報探索内容それぞれの利用度は以下のグラフ 23 の通りである。今回追加した項目は「オンラインの就職ガイダ

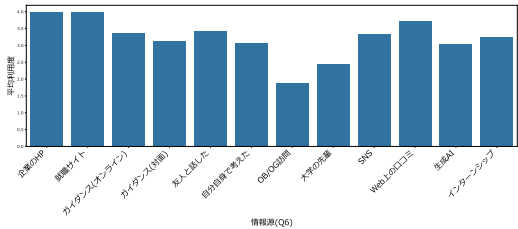


図 2 情報源の平均利用度

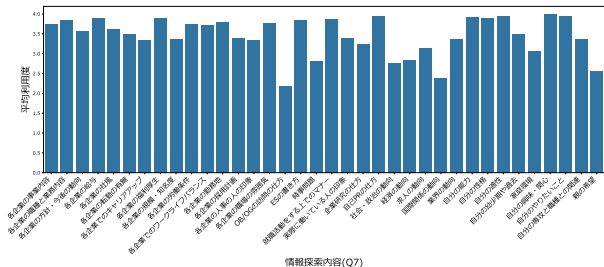


図 3 情報探索内容の平均利用度

ンスに参加した」「OB/OG 訪問をした」「SNS で就職活動に関する情報を閲覧・検索した」「WEB 上で就職活動に関する口コミ情報を閲覧・検索した」「生成 AI と就職活動について話した」「インターンシップに参加した」の 6 項目であるが、いずれも従来の項目と同等の平均利用度を示しており、十分に利用されていることがわかる。

4.1.3 情報探索内容ごとの情報源の利用度

本調査では、情報源と情報探索内容の利用の関係性についても調べるため、情報探索内容ごとの情報源の利用度に関する設問を設定した。各情報探索内容について、「全く行っていない」以外を選択した被験者に、情報源に関する設問と同等の内容から、その情報探索内容を調べる際にどの情報源を利用したかを複数選択で尋ねた。以下の図 4 は、各情報探索内容について、「全く行っていない」以外を選択した分析対象者の内、その情報探索内容を調べる際に各情報源を利用したと答えた分析対象者の割合を示している。この図を参考に、因子分析を行う。

4.2 単変数間の相関分析

始めに、行動経済学的個人特性と性別、就職活動結果(就職活動自体に対する満足度・入社予定企業に対する満足度)の相関を調べた。結果をグラフ 5 に示す。

時間割引率・現在バイアスと就職活動の結果の相関に有意差は見られなかった。一方、リスク選好と就職活動に関しては相関に有意差が見られ、リスク愛好的な学生は就職活動に対する満足度がやや下がるという傾向が見られた。また、就職活動自体の満足度と入社予定企業への満足度は強い正の相関を表した。

次に、行動経済学的個人特性と情報源・情報探索内容の利用度の相関を調べた。全体として強い相関は見られなかったものの、以下の表 7 に示すような中程度の相関関係に有意差が見られた。

時間割引率に関しては、就職サイトの利用度とインターン

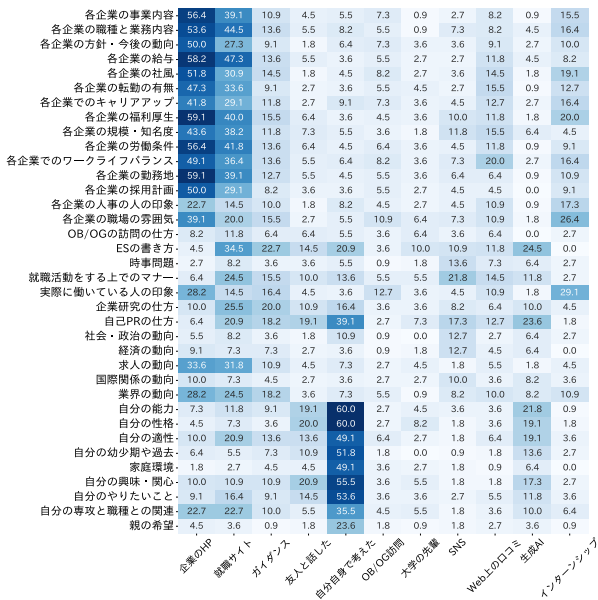


図 4 情報探索内容ごとの情報源の利用度

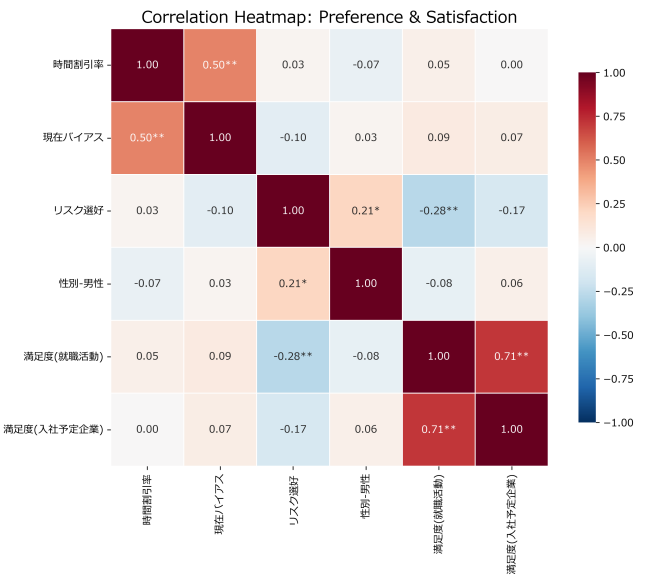


図 5 各選好・性別・就職活動結果の相関分析

表 7 行動経済学的個人特性と情報源の相関分析 (p < 0.05)

選好	情報源	相関係数	p
時間割引率	就職サイト	-0.235	0.0133
時間割引率	インターンシップ	0.190	0.0467
現在バイアス	友人と話した	-0.218	0.0223
現在バイアス	生成 AI	-0.242	0.0110
リスク選好	生成 AI	0.238	0.0123

シップの参加度との相関が見られ、時間割引率が高い、すなわちせっかちな学生の方がインターンシップの参加度がやや高く、就職サイトの利用率が低い傾向が見られた。現在バイアスに関しては友人と話す、生成 AI と話すといった情報探索行動との相関が見られ、いずれも現在バイアスが強い、先延ばし傾向の強い学生の方がこれら利用率が低い傾向が見られた。また、リスク選好に関しても生成 AI の利用度に差が見られ、リスク

表 8 行動経済学的個人特性と情報探索内容の相関分析 ($p < 0.05$)

選好	情報探索内容	相関係数	p
時間割引率	各企業でのキャリアアップ	-0.249	0.0088
時間割引率	自分の幼少期や過去	-0.211	0.0270
時間割引率	家庭環境	-0.213	0.0251
現在バイアス	各企業の福利厚生	-0.196	0.0404
現在バイアス	各企業でのワークライフバランス	-0.206	0.0309
現在バイアス	自己 PR の仕方	-0.237	0.0126
現在バイアス	自分の適性	-0.239	0.0119
現在バイアス	自分の幼少期や過去	-0.359	0.0001
現在バイアス	家庭環境	-0.265	0.0052
リスク選好	自分の性格	-0.190	0.0465
リスク選好	自分の興味・関心	-0.214	0.0245

表 9 就職活動結果と情報源の相関分析 ($p < 0.05$)

結果	情報源	相関係数	p
満足度 (就職活動)	企業の HP	0.235	0.0134
満足度 (就職活動)	友人と話した	0.211	0.0272
満足度 (就職活動)	インターンシップ	0.349	0.0002
満足度 (入社予定企業)	企業の HP	0.198	0.0376

選好が強く、リスク愛好的な学生ほど生成 AI をよく使用する傾向が見られた。

情報探索内容と行動経済学的個人特性については、次のような中程度の相関関係に有意差が見られた。まず、時間割引率は「各企業でのキャリアアップ」「自分の幼少期や過去」「家庭環境」の探索度合いとの相関を示し、時間割引率が高く、せっかちな学生の方がこれらの情報の探索度合いが低くなる傾向が見られた。次に現在バイアス・リスク選好についても複数の情報の探索度合いとの相関が見られたが、いずれも負の相関を示し、現在バイアスが強い、すなわち先延ばし傾向の強い学生ほど「各企業の福利厚生」「各企業でのワークライフバランス」「自己 PR の仕方」「自分の適性」「自分の幼少期や過去」「家庭環境」といった情報探索の度合いが低くなり、リスク選好ではよりリスク愛好的な学生の方が「自分の性格について考える」「自分の興味・関心について考える」といった情報の探索度合いが低くなる傾向が見られた。

同様に、情報源・情報探索内容の利用度と就職活動結果との相関を調べた。相関係数、 p 値を表 9、表 10 に示す。

情報源のうち、就職活動自体の満足度と正の相関が見られたのは「企業の HP」「友人」「インターンシップ」であった。また、入社予定企業に対する満足度と正の相関が見られた情報源は「企業の HP」であった。

情報探索内容については、「各企業の方針・今後の動向」といった企業に関する情報、「自己 PR の仕方」や「業界の動向」といった就職活動全体に関する情報、「自分の能力」「自分の性格」「自分の適性」「自分の幼少期や過去」「自分の興味・関心」「自分のやりたいこと」といった自分自身に関する情報で、就職活動自体の満足度との正の相関に有意差が見られた。また、「自分の専攻と職種との関連」についても就職活動自体の満足度と正の相関が見られた。一方、「親の希望を考える」という情報探索においてのみ、就職活動自体の満足度と負の相関が見

表 10 就職活動結果と情報探索内容の相関分析 ($p < 0.05$)

結果	情報源	相関係数	p
満足度 (就職活動)	各企業の方針・今後の動向	0.233	0.0142
満足度 (就職活動)	自己 PR の仕方	0.232	0.0148
満足度 (就職活動)	業界の動向	0.204	0.0327
満足度 (就職活動)	自分の能力	0.331	0.0004
満足度 (就職活動)	自分の性格	0.297	0.0016
満足度 (就職活動)	自分の適性	0.286	0.0024
満足度 (就職活動)	自分の幼少期や過去	0.189	0.0477
満足度 (就職活動)	自分の興味・関心	0.331	0.0004
満足度 (就職活動)	自分のやりたいこと	0.255	0.0072
満足度 (就職活動)	自分の専攻と職種との関連	0.198	0.0377
満足度 (就職活動)	親の希望	-0.296	0.0017
満足度 (入社予定企業)	自己 PR の仕方	0.222	0.0198
満足度 (入社予定企業)	自分の能力	0.247	0.0092
満足度 (入社予定企業)	自分の性格	0.269	0.0044
満足度 (入社予定企業)	自分の適性	0.272	0.0040
満足度 (入社予定企業)	自分の興味・関心	0.318	0.0007
満足度 (入社予定企業)	自分のやりたいこと	0.270	0.0044
満足度 (入社予定企業)	親の希望	-0.305	0.0012

れた。入社予定企業に対する満足度についても「自己 PR の仕方」といった就職活動全体に関する情報、「自分の能力について考えた」「自分の性格について考えた」「自分の適性について考えた」「自分の興味・関心について考えた」「自分のやりたいことについて考えた」といった自分自身に関する情報において正の相関が見られた。また、就職活動自体に対する満足度と同様に「親の希望を考える」という情報探索においてのみ、負の相関が見られた。

4.3 情報探索行動の因子分析

以上から、情報探索行動と複数の項目間において中程度の相関が認められた。そこで、これらの項目間に共通する潜在的な構造を明らかにし、変数を集約してその後の分析に用いることを目的とし、情報源・情報探索内容それぞれについて因子分析を行った。因子の抽出法には最尤法を用い、因子間の相関を仮定してプロマックス回転を適用した。因子数の決定にあたっては、カイザー基準 (固有値 1 以上) およびスクリープロットの形状をもとに判断した。

4.3.1 情報源の因子分析

情報源の因子負荷量を表 11 に示す。情報源に関する 12 項目について、スクリープロットに基づいて 5 因子構造を採用した。この時、累積寄与率は 52.4%であった。また、因子負荷量 0.40 以上を基準として各項目に着目すると、第 1 因子は「企業の HP」「就職サイト」「ガイダンス (オンライン)」「SNS」「Web 上の口コミ」、第 2 因子は「ガイダンス (対面)」、第 3 因子は「友人」「大学の先輩」「インターンシップ」、第 4 因子は「OB/OG 訪問」、第 5 因子は「生成 AI」の因子負荷量が高くなっていることがわかる。また、「自分自身で考えた」については、いずれの因子にも十分な負荷量を示しておらず、かつ共通性も 0.06 と低くなっている。

情報探索行動に関しても、情報源と同様にスクリープロットに基づいて因子数を決定し、7 因子構造を採用した。この時、

表 11 情報源に関する探索的因子分析の結果（最尤法、プロマックス回転）。F1～F5 は因子番号を表す。

	F1	F2	F3	F4	F5	h^2
hp	0.73	-0.04	0.16	-0.13	-0.07	0.58
site	0.92	0.05	-0.18	-0.03	-0.18	0.91
guide_online	0.42	0.18	-0.23	0.38	0.17	0.43
guide_real	-0.17	1.00	0.08	0.17	0.16	1.08
friend	0.03	0.13	0.53	-0.11	0.12	0.32
self	0.07	0.21	0.04	-0.05	-0.11	0.06
obog_visit	-0.14	0.04	0.19	0.75	-0.09	0.63
senior	-0.14	0.05	0.63	0.23	0.01	0.47
sns	0.50	-0.02	0.25	-0.06	0.28	0.40
review	0.75	-0.08	0.00	0.02	0.10	0.58
ai	-0.02	-0.02	0.07	-0.08	0.68	0.47
internship	0.19	-0.14	0.41	0.32	-0.13	0.34
SS Loadings	2.47	1.12	1.07	0.94	0.69	
Proportion Var	0.21	0.09	0.09	0.08	0.06	
Cumulative Var	0.21	0.30	0.39	0.47	0.52	

累積寄与率は 55.9%であった。各項目についても情報源と同様に因子負荷量 0.40 以上を基準として解釈すると、第 1 因子は「各企業の事業内容」「各企業の職種と業務内容」「各企業の給与」「各企業の社風」「各企業の転勤の有無」「各企業の福利厚生」「各企業の規模・知名度」「各企業の労働条件」「各企業のワークライフバランス」「各企業の勤務地」「各企業の職場の雰囲気」、第 2 因子は「ES の書き方」「就職活動をする上でのマナー」「自己 PR の仕方」「自分の能力」「自分の性格」「自分の適性」「自分の興味・関心」「自分のやりたいこと」「自分の専攻と職種との関連」、第 3 因子は「OB/OG 訪問の仕方」「時事問題」「社会・政治の動向」「経済の動向」「国際関係の動向」、第 4 因子は「実際に働いている人の印象」「業界の動向」「親の希望」、第 5 因子は「求人の動向」、第 6 因子は「各企業の転勤の有無」「各企業の勤務地」「家庭環境を振り返った」、第 7 因子は「各企業の事業内容」となった。また、「各企業の方針・今後の動向」「各企業でのキャリアアップ」「各企業の採用計画」「各企業の人事の人の印象」「企業研究の仕方」「自分の幼少期や過去」については、いずれの因子にも十分な負荷量を示さなかった。「各企業の事業内容」については 2 つの因子に 0.40 以上の負荷を示した。

4.4 重回帰分析

次にこれらの抽出された各因子が満足度に対してどのような影響を及ぼしているのかを明らかにするため、各因子を説明変数とした重回帰分析を行う。従属変数は就職活動の結果（就職活動に対する満足度・入社予定企業への満足度）である。また、因子分析の結果、情報源の「自分自身で考えた」の共通性（ h^2 ）は 0.06 と著しく低く、抽出された 5 つの因子のいずれに対しても十分な負荷量を示さなかった。これは、当該項目が他の外部情報源（メディアや対人ネットワーク）とは性質が根本的に異なる、独自の概念（内省的思考や自己決定プロセス）であることを示唆している。そのため、重回帰分析においては、情報源の因子分析によって抽出された 5 因子（因子得点）と情報探索内容の因子分析によって抽出された 7 因子に加え、この

表 12 情報探索内容に関する因子分析結果（最尤法、プロマックス回転）

	F1	F2	F3	F4	F5	F6	F7	h^2
bc	0.50	0.10	-0.13	-0.07	-0.02	0.05	0.49	0.52
occupation	0.70	0.13	-0.16	-0.04	0.03	0.03	0.03	0.54
future	0.17	0.26	0.09	0.23	0.13	-0.01	0.11	0.19
salary	0.87	-0.16	-0.05	-0.06	0.10	-0.01	-0.05	0.81
culture	0.70	-0.06	-0.05	0.03	0.07	0.16	0.18	0.56
transfer	0.45	-0.05	-0.00	0.22	-0.02	0.42	0.02	0.42
career	0.07	0.18	0.07	0.28	0.29	0.18	0.00	0.23
benefit	0.84	-0.02	-0.00	-0.10	0.06	0.05	0.02	0.72
large	0.61	-0.11	0.24	-0.02	-0.02	-0.12	0.16	0.49
labor_condition	0.74	0.09	-0.13	0.07	-0.04	0.04	0.06	0.58
WLB	0.57	0.05	0.01	0.05	0.02	0.07	0.22	0.39
place	0.73	0.02	0.10	-0.24	-0.15	0.52	0.00	0.89
ricruit_plan	0.18	0.03	0.31	-0.13	0.24	0.13	-0.09	0.23
HR	0.31	-0.03	0.19	0.24	-0.15	0.01	0.09	0.22
atmosphere	0.63	0.11	-0.02	0.24	-0.09	0.01	-0.12	0.49
OBOG	-0.13	-0.11	0.48	0.28	-0.14	-0.07	0.13	0.38
ES	0.38	0.41	0.11	-0.33	-0.03	-0.23	0.24	0.54
news	0.00	0.09	0.88	-0.14	-0.14	0.19	0.05	0.87
manner	0.28	0.58	0.06	0.08	0.01	-0.12	-0.39	0.59
impression	0.35	0.11	-0.03	0.66	-0.24	-0.05	-0.12	0.65
search_company	0.23	-0.11	0.21	0.19	0.16	0.01	0.18	0.20
selfPR	0.02	0.73	-0.18	-0.10	0.22	0.02	0.08	0.63
society	-0.00	0.01	0.82	-0.04	0.09	0.03	-0.11	0.69
economics	0.02	-0.06	0.81	0.12	0.08	0.02	-0.28	0.77
hiring_trends	0.02	-0.17	-0.07	0.00	0.90	0.01	-0.00	0.84
international	-0.16	0.02	0.79	0.13	-0.09	-0.02	-0.01	0.67
industry	0.27	-0.05	0.03	0.47	0.30	-0.05	-0.11	0.40
capability	0.04	0.70	-0.03	0.03	0.13	0.02	-0.05	0.52
personality	0.05	0.88	-0.08	0.11	-0.19	0.10	-0.10	0.85
aptitude	0.05	0.82	-0.05	0.03	-0.08	-0.01	0.21	0.73
childhood	0.02	0.39	-0.12	0.28	0.16	0.13	0.21	0.33
family	-0.22	0.27	0.12	0.23	0.15	0.50	0.11	0.47
interest	-0.13	1.05	0.01	0.10	-0.25	0.09	-0.11	1.20
want	-0.10	0.94	0.11	-0.16	-0.20	0.18	0.02	1.00
major	-0.05	0.46	0.18	-0.06	0.19	-0.18	-0.06	0.32
paren	-0.13	-0.06	0.08	0.41	0.06	0.08	-0.02	0.21
SS Loadings	5.86	5.67	3.42	1.70	1.57	1.01	0.91	
Proportion Var	0.16	0.16	0.10	0.05	0.04	0.03	0.03	
Cumulative Var	0.16	0.32	0.42	0.46	0.51	0.53	0.56	

「自分自身で考えた」を独立した 1 変数として投入することとした。

結果を以下の表 13、14 に示す。

就職活動に対する満足度を従属変数とした重回帰分析では、決定係数（ R^2 ）は 0.266、自由度調整済み決定係数は 0.140 であった。分散分析の結果、回帰モデル全体として有意な適合が認められた（ $F(16, 93)=2.108, p=0.014$ ）。また、各独立変数が従属変数に与える影響を検討したところ、情報源の第 3 因子（「友人」「大学の先輩」「インターンシップ」）、情報探索内容の第 2 因子（「ES の書き方」「就職活動をする上でのマナー」「自己 PR の仕方」「自分の能力」「自分の性格」「自分の適性」「自分の興味・関心」「自分のやりたいこと」「自分の専攻と職種との関連」）が就職活動に対する満足度に正の有意な影響を及ぼしていた。

入社予定企業に対する満足度を従属変数とする重回帰分析で

表 13 重回帰分析 (従属変数: 就職活動に対する満足度)

	回帰係数	標準誤差	p
const	5.123	0.422	0.000
Src_F1	0.137	0.188	0.468
Src_F2	0.004	0.118	0.974
Src_F3	0.412	0.160	0.012
Src_F4	0.095	0.155	0.543
Src_F5	-0.242	0.151	0.113
Cont_F1	-0.161	0.187	0.391
Cont_F2	0.367	0.153	0.018
Cont_F3	-0.158	0.148	0.289
Cont_F4	-0.143	0.168	0.398
Cont_F5	0.042	0.159	0.790
Cont_F6	-0.141	0.116	0.227
Cont_F7	-0.006	0.149	0.969
switch-1	-0.010	0.054	0.849
p-bias	0.189	0.197	0.341
risk_switch	-0.091	0.054	0.097
Q6-6	-0.086	0.091	0.342

調整済み R^2 : 0.14

表 14 重回帰分析 (従属変数: 入社予定企業に対する満足度)

	回帰係数	標準誤差	p
const	4.800	0.335	0.000
Src_F1	-0.106	0.161	0.513
Src_F3	0.338	0.149	0.026
Src_F5	-0.180	0.134	0.181
Cont_F1	-0.132	0.170	0.436
Cont_F2	0.421	0.131	0.002
Cont_F3	-0.068	0.122	0.578
Cont_F4	-0.159	0.150	0.289
Cont_F6	-0.063	0.109	0.563
Cont_F7	0.062	0.131	0.636
switch-1	-0.050	0.050	0.316
p-bias	0.254	0.177	0.154
Q6-6	-0.040	0.084	0.633

調整済み R^2 : 0.09

は同数の独立変数では回帰モデル全体の有意性が確保されなかったため ($p=0.121$), ステップワイズ変数減少法に基づき, 有意確率 (p 値) の高い変数から順に除外する検討を行った. 最終的に, 表に示す 12 変数によってモデル全体の有意性が確保された. 修正後のこのモデルでは決定係数 (R^2) は 0.193, 自由度調整済み決定係数は 0.093 であった. 分散分析の結果, F 値は 1.930、有意確率は 0.0397 であった. また, 各独立変数が従属変数に与える影響を検討したところ, 就職活動に対する満足度と同じく, 情報源の第 3 因子 (友人)「大学の先輩」「インターンシップ」と情報探索内容の第 2 因子 (「ES の書き方」「就職活動をする上でのマナー」「自己 PR の仕方」「自分の能力」「自分の性格」「自分の適性」「自分の興味・関心」「自分のやりたいこと」「自分の専攻と職種との関連」) が就職活動に対する満足度に正の有意な影響を及ぼしていた.

5 考 察

5.1 分析結果の解釈

5.1.1 尺度の検討

本研究では, 矢崎ら [19] の「就職における情報探索行動尺度」を参考に, 項目の変更・設問方法の変更を行い, 質問紙調査を実施した. その結果, 本研究で追加した「オンラインの就職ガイダンスに参加した」「OB/OG 訪問をした」「SNS で就職活動に関する情報を閲覧・検索した」「WEB 上で就職活動に関する口コミ情報を閲覧・検索した」「生成 AI と就職活動について話した」「インターンシップに参加した」のいずれの項目も高い利用度を示しており, 現代の就職活動における情報探索行動をより正確に捉えることができた. これにより, 情報探索行動の研究において, 日々変化していく情報利用の実態を逐次に更新していくことの重要性が明らかとなった. また, 設問方法に関しても, 図 4 に示すように, 情報源と情報探索内容は多対多の関係性で複雑な組み合わせを形成していることがわかった. 調査者によって独自に組み合わせるのではなく, 情報源と情報探索内容の利用度合いを独立に調査し, 加えてその関係性について調査することでより詳細に情報探索行動を捉え, かつ調査対象者の回答の負担を減らすことができた.

5.1.2 情報探索行動の分析

始めに, 先述の情報源と情報探索内容の関係性について考察する. 「各企業の事業内容」をはじめ, 「給与」, 「キャリアアップ」, 「福利厚生」など, 各企業の情報については「企業の HP」が最も使用されていた. また, 同様の内容をマイナビ・リクナビといった就職サイトを用いて探索している学生も多く見られた. ただし, 各企業情報に加えて, 就職サイトは「ES の書き方」「企業研究の仕方」「専攻と職種との関連」といった広く就職活動全般に関する情報の探索にもよく利用されており, 汎用的な情報を含む情報源であることがわかった. 次に特徴的であったのは「自分の性格について考えた」「自分の能力について考えた」「自分のやりたいことについて考えた」といった自分自身に関する情報である. 情報源として「自分自身」が多く選択されていたのは言わずもがなであるが, 次いで「生成 AI と対話した」や「友人と話した」が多く選択されており, 一人で考えるだけでなく, 対話することによって, 自分自身に関する理解を深めていることがわかった. また, 同データを情報源ごとに考察すると, 「OB/OG 訪問をした」「大学の先輩に聞いた」「SNS で閲覧・検索した」の 3 つの情報源はいずれの情報探索内容においてもあまり用いられていないことがわかる. OB/OG 訪問や大学の先輩は社会人との個人的な交流に基づいた情報探索行動であり, 「情報源」としての働きではなく, コネクションの形成のような別の働きを持っている可能性が考えられる. SNS に関しては, モバイル社会研究所による調査を参照して考察したい [10]. 20 代においてよく利用されている SNS である X や Instagram は利用目的において「暇つぶしに使う」がそれぞれ 2 位, 1 位となっており, そもそも情報源として明確な意図を持って使用されることが少ないことがわかる. そのため「情報

探索」ではなく「SNS を閲覧する」ことが目的化し、得られる情報に一貫した特徴が見られないと考えられる。また、「口コミ」や「インターンシップ」については利用率はやや下がるものの、HP や就職サイトと同様に各企業の情報を探索する際によく使われており、前者のような公的情報の補足として、実際に働いている人から企業に関する情報を取得しようとしていることが読み取れる。

5.1.3 行動経済学的個人特性と情報探索行動の関連

次に、RQ1, RQ2 を明らかにするため、行動経済学的個人特性と情報探索行動の関連について考察する。

情報源：情報源に関しては時間割引率と就職サイト・インターンシップの利用度合いに相関が見られた。一般的に、インターネットを通して手軽に閲覧できる就職サイトに比べて、複数回の選考が必要となるインターンシップは、情報源として利用するために選考を乗り越える忍耐強さが必要なため、時間割引率が低い(我慢強い)学生の方がインターンシップの参加率が高いことが予想される。しかし、実際の結果では、時間割引率が高く、せっかちな学生の方がインターンシップの参加度が高く、就職サイトの利用率が低いことがわかった。この傾向や先述の情報探索内容との関連を考慮すると、時間割引率が低い学生は就職活動における最終的な報酬である内定を見据え、就職サイトのような公的かつ幅広い情報にアクセスしているのに対し、時間割引率が高くせっかちな学生は「インターンシップへの参加」を目先の報酬と捉え、就職サイトのような汎用的な情報源へのアクセスが少なくなっている可能性が考えられる。次に、現在バイアスについては先延ばし傾向の強い学生の方が友人や生成 AI との対話が少ない傾向が見られた。生成 AI についてはリスク選好との相関も見られ、リスク愛好的な学生の方が多く使用していた。図 4 において友人・生成 AI との対話内容の多くが自分自身に関する情報であったことを鑑みると、より未来志向的な学生の方が、自分について客観的に捉えようとしていることが考えられる。また、リスク選好と生成 AI の利用度の関係は Chatterjee, Eliashberg の研究によって説明できる [2]。この研究は、新たな製品やサービスの普及率についての Rogers のイノベーション普及理論 (Diffusion of Innovations) [6] に、リスク選好などに代表される個々人の意思決定の行動原理を組み込んだものである。この Chatterjee らのモデルにおいて、リスク愛好的な人は「即時採用者 (immediate adopters)」と呼ばれるタイプに分類され、新たな技術やサービスが生み出された際に即座にそれらを採用する消費者であるとされている。本研究においても、近年急速な成長を見せている生成 AI に対してリスク愛好的な学生の利用率が高い傾向が見られ、これは先述のような先行研究と合致する結果である。

情報探索内容：次に情報探索内容の利用度と行動経済学的個人特性との関係である。まず、全体として負の相関が見られた点に着目したい。時間割引率に代表されるせっかちなさ、現在バイアスに代表される先延ばし傾向、リスク選好に代表されるリスク愛好度のいずれにおいても、突出した情報探索が行われるのではなく、特定の情報探索が減少することが明らかになった。特に、時間割引率では「自分の幼少期や過去」「家庭環境」、現

在バイアスでは、同じく「自分の幼少期や過去」「家庭環境」に加えて「自分の適性」、リスク選好では「自分の性格」「自分の興味・関心」の項目の探索度合いが低くなっており、それぞれに代表される性質が強くなるにつれ、自分自身に関する探索が減少してしまう傾向が見られた。これらのことからせっかちな学生や先延ばし傾向の強い学生においては内省を必要とする自己分析がコストとして捉えられていることが伺える。またリスク愛好的な学生においては、将来的なリスクである、自己に関する探索不足による自分自身と企業とのミスマッチが軽んじられていることが考えられる。

その他にはせっかちな学生ほど「各企業でのキャリアアップ」に関する探索度合いが低い、先延ばし傾向の強い学生ほど「各企業での福利厚生」や「各企業でのワークライフバランス」に関する探索度合いが低いといった傾向が見られ、入社後の持続的な環境に対する情報探索が少なくなっていることが明らかになった。

5.1.4 情報探索行動と就職活動結果

前セクションでは RQ1, 2 に従い、行動経済学的個人特性と情報探索行動の関連について考察した。本セクションでは RQ3 を明らかにするため、情報探索行動と就職活動結果の分析結果に対する考察を述べる。

単変数間での分析：まずは 12 項目の情報源、36 項目の情報探索内容と 2 種類の就職活動結果 (就職活動に対する満足度・入社予定企業に対する満足度) の相関分析について述べる。

情報源のうち、就職活動結果と相関を示した項目は「企業の HP」「友人と話した」「インターンシップ」であり、いずれも正の相関が見られた。特にインターンシップは就職活動の満足度と強い正の相関を示した。また、この 4 つの内「企業の HP」のみが双方の満足度 (就職活動に対する満足度・入社企業に対する満足度) と正の相関を示した。「友人と話した」について、就職活動不安とソーシャル・サポートについて調査した田原らの研究では、友人からの情動的及び情緒的サポートと就職活動不安との間に有意な関連が見られなかったことから [16]、友人とのコミュニケーションは精神的なアプローチを通して就職活動の満足度を高めているのではなく、情報探索をサポートする役割として就職活動の満足度に寄与している可能性が考えられる。

情報探索内容については、就職活動に対する満足度では「自分の能力」「自分の性格」「自分の適性」「自分の幼少期や過去」「自分の興味・関心」「自分のやりたいこと」「自分の専攻と職種との関連」、入社企業に対する満足度では「自分の能力」「自分の性格」「自分の適性」「自分の興味・関心」「自分のやりたいこと」の自分自身に関する探索を示す項目がいずれも中程度の有意な正の相関を示した。一方で、いずれの満足度とも負の相関を示した唯一の情報探索内容としては「親の希望」が挙げられ、対照的な結果が得られた。また、双方の満足度に正の相関を示した情報探索内容には「自己 PR の仕方」が見られた。これらの結果から、就職活動における満足度を高めるためには自分自身を深く内省的に見つめるとともに、その伝達まで考慮することが重要であることが考えられる。

重回帰分析を用いた分析：最後に、これらの分析結果を包括

的に考察するために、情報源・情報探索内容それぞれに対して因子分析を行い、さらに2つの満足度を従属変数とした重回帰分析を行った。重回帰分析では因子分析によって得られた5つの情報源因子と7つの情報探索内容因子、3つの行動経済学的個人特性を独立変数とした。双方の満足度において、自由度調整済み決定係数の値は低かったものの、同様の独立変数、情報源第3因子と情報探索内容第2因子が有意な相関を示し、決定係数も近い値であった。情報源第3因子は「友人と話した」「大学の先輩に聞いた」「インターンシップに参加した」といった対人関係に基づいた内容であり、情報探索内容第2因子は「ESの書き方」「就職活動をする上でのマナー」「自分の能力」「自分の性格」「自分の適性」「自分の幼少期や過去」「自分の興味・関心」「自分のやりたいこと」「自分の専攻と職種との関連」といった自分自身に関する内省的な情報であった。これらは単変数間での分析と一致するものであり、コミュニケーションを通して得られる情報にアクセスすること、自分自身について探索することが就職活動をする上で重要となっていることがわかる。また、本章で述べた情報探索行動の分析において、自分自身についての情報を探索する際に利用する情報源として、「自分自身で考えた」について「友人と話した」が選択されていたことや重回帰分析において加えた「自分自身で考えた」の項目が有意な相関を示さなかったことを考えると、自己に関する情報を他者からの客観的な視点に基づいて理解することが非常に重要であると言える。

5.2 介入の検討

RQ3について、より具体的な支援として、行動経済学で用いられる「ナッジ」について検討するため、構造方程式モデリングを用いた分析を行う。特に、本研究ではこれまでの分析で就職活動や行動経済学的個人特性と強い相関を示した自分自身に関する情報探索行動に着目し、因果モデルを構築する。検証する因果モデルはこれまでの考察を踏まえ、次のように定めた。まず、最終的な従属変数はこれまでと同様に就職活動の結果(就職活動に対する満足度・入社予定企業に対する満足度)とする。次に、独立変数として、因子分析で求めた情報源第3因子(「友人と話した」「大学の先輩に聞いた」「インターンシップに参加した」)、情報探索内容第2因子(「ESの書き方」「就職活動をする上でのマナー」「自分の能力」「自分の性格」「自分の適性」「自分の幼少期や過去」「自分の興味・関心」「自分のやりたいこと」「自分の専攻と職種との関連」)、3つの行動経済学的個人特性を設定した。モデルは行動経済学的個人特性が情報源・情報探索内容・結果の順に影響を及ぼすもの(図6)、行動経済学的個人特性が情報探索内容・情報源・結果の順に影響を及ぼすもの(図7)、情報源と情報探索内容が相関関係を形成するもの(図8)、情報源と情報探索内容が独立して存在するもの(図9)の4つのモデルを作成した。検証結果のパス図と各モデルの適合後を表すCFIを以下に示す。

いずれもCFIは0.5前後であり、不適合となった。相関分析において見られた他の情報源・情報探索内容との関係性の考慮や、就職活動における情報探索行動以外の選考などの要素の考



図6 モデル1(CFI:0.512)

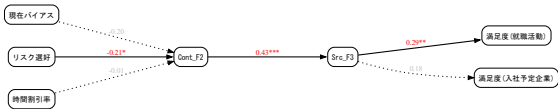


図7 モデル2(CFI:0.486)

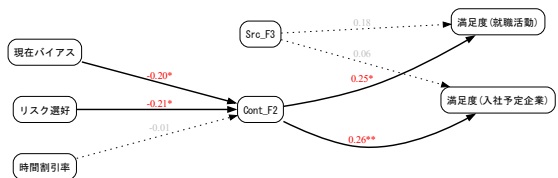


図8 モデル3(CFI:0.580)

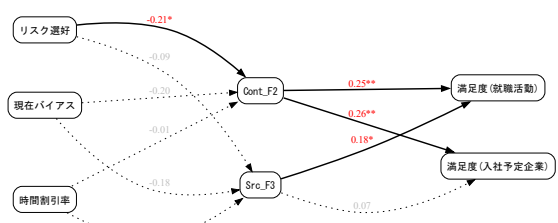


図9 モデル4(CFI:0.417)

慮が必要である。一方で、一部のモデルでは、リスク選択や現在バイアスと情報探索内容、また情報探索内容と満足度との因果関係が見られた。そのため、これらの選好を測定し、今回変数として導入した自己探索に代表される情報探索を促すことは介入策としての効果が期待される。

5.3 研究の限界

最後に本研究の限界と今後の展望について述べる。第1に、サンプルの偏りと規模の問題である。分析対象者は110名と限定的であり、その内訳も女性が85名と過半数を占めているため、結果を一般化するにはさらなるデータの蓄積が必要である。第2に、本研究は就職活動を終えた時点での調査である点である。本来、長期化している就職活動はインターンシップを受ける時期や、本選考の時期など、時期ごとに異なる情報探索の必要性がある[18]。しかし、本調査では就職活動終了時点での累積的な情報探索行動しか捉えられていない。今後、より実際の支援に寄与する研究を行っていくには、就職活動の時系列に沿った情報探索行動やその個人差を明らかにする必要があると考えられる。第3に、就職活動の情報探索行動以外の点を

捉えられていない点である。本研究は情報探索行動と行動経済学的個人特性に焦点を絞ってその関係性を調査した。しかし、就職活動は複数回にわたる選考など、情報探索行動以外の要素も多分に含むものである。そのため、より包括的な就職活動のモデルとその個人差を分析することが重要である。最後に、就職後の満足度を捉えられていない点である。就職活動におけるゴールは内定の獲得であるが、その後にも早期離職などの課題は存在する。より長い視点で「より良い就職活動とは何か」を考えるため、縦断調査を用いた長期的なデータの収集・分析が必要であると考えられる。

6 おわりに

本研究は、新規学卒者の就職活動が早期化・長期化する現状において、学生個々人の特性に応じた実践的な支援の在り方を検討することを目的に、行動経済学的な個人特性（時間選好・リスク選好）が情報探索行動および就職活動の満足度に与える影響を分析した。本研究を通じて得られた主要な結論は以下の通りである。

第1に、現代の就職活動における情報源および探索内容の実態を捉えるために既存の尺度を更新し、その有効性を確認した。新たに追加した生成 AI や SNS、オンラインガイダンスといった項目は、従来の企業の HP や就職サイトと同等の高い平均利用度を示しており、情報の多角化が進む現代の状況をより正確に反映した分析が可能となった。

第2に、行動経済学的特性と情報探索行動との間には有意な負の相関が認められ、特定の特性が探索の「偏り」を規定していることが明らかとなった。具体的には、現在バイアスが強く「先延ばし傾向」のある学生ほど、自己分析に関連する探索（自己 PR、適性、幼少期の振り返り等）の度合いが著しく低い（現在バイアスと自分の幼少期や過去の相関係数： -0.359 ）といった傾向が見られた。また、時間選好が高い「せっかち」な学生ほどキャリアアップに関する探索を、先延ばし傾向の強い学生ほど福利厚生やワークライフバランスといった長期的な就業継続に関わる情報の探索を避ける実態も示された。

第3に、就職活動の満足度を左右する決定要因を特定した。重回帰分析の結果、対人関係や実体験に基づく「情報源第3因子（友人、先輩、インターンシップ）」および、自己理解を深める「情報探索内容第2因子（自己分析関連）」が、就職活動自体および入社予定企業への満足度に対して正の有意な影響を及ぼしていた。構造方程式モデリング（SEM）による検証（モデル3： $CFI = 0.580$ ）からも、行動特性が自己探索行動を媒介して満足度に影響を与える因果構造の可能性が示唆された。以上の知見は、就職活動の効率化や質の向上を支援する上で、一律の情報提供ではなく、学生の意思決定特性に着目した介入（ナッジ）の必要性を示唆している。特に内省を「コスト」と感じやすい現在志向的な学生に対しては、生成 AI や友人との対話といった、客観的視点を容易に得られる手段を初期設定として組み込むなどの支援が有効であると考えられる。

今後の課題として、本研究は110名という限定的なサンプル

サイズかつ女性に偏った属性での分析にとどまっており、結果の一般化にはさらなるデータの蓄積が必要である。また、今後は就職活動の時系列に沿った縦断的調査や、入社後の満足度までを含めた長期的な影響評価を行うことで、より包括的な意思決定支援モデルへと発展させていくことが望まれる。

謝 辞

本研究の一部は、JSPS 科研費 24K03046、24K23877 の助成を受けたものです。ここに記して謝意を示します。

文 献

- [1] George Ainslie. Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological bulletin*, Vol. 82, No. 4, p. 463, 1975.
- [2] Rabik Ar Chatterjee and Jehoshua Eliashberg. The innovation diffusion process in a heterogeneous population: A micro modeling approach. *Management science*, Vol. 36, No. 9, pp. 1057–1079, 1990.
- [3] Maribeth Coller and Melonie B Williams. Eliciting individual discount rates. *Experimental economics*, Vol. 2, No. 2, pp. 107–127, 1999.
- [4] Charles A Holt and Susan K Laury. Risk aversion and incentive effects. *American economic review*, Vol. 92, No. 5, pp. 1644–1655, 2002.
- [5] Kris N Kirby and Nino N Maraković. Delay-discounting probabilistic rewards: Rates decrease as amounts increase. *Psychonomic bulletin & review*, Vol. 3, No. 1, pp. 100–104, 1996.
- [6] Everett M Rogers, Arvind Singhal, and Margaret M Quinlan. Diffusion of innovations. In *An integrated approach to communication theory and research*, pp. 432–448. Routledge, 2014.
- [7] Stephen A Stumpf, Stephen M Colarelli, and Karen Hartman. Development of the career exploration survey (ces). *Journal of vocational behavior*, Vol. 22, No. 2, pp. 191–226, 1983.
- [8] Richard H. Thaler and Cass R. Sunstein. *Nudge : improving decisions about health, wealth and happiness*. Penguin books, revised ed edition, 2009.
- [9] Zhuoran Zhang, Xiang Yao, Shuai Yuan, Yingxin Deng, and Chao Guo. Big five personality influences trajectories of information seeking behavior. *Personality and Individual Differences*, Vol. 173, p. 110631, 2021.
- [10] NTT ドコモモバイル社会研究所. モバイル社会白書 2025 年版. White paper, NTT ドコモ モバイル社会研究所, 2025.
- [11] 下村英雄, 堀洋元. 大学生の就職活動における情報探索行動: 情報源の影響に関する検討. *社会心理学研究*, Vol. 20, No. 2, pp. 93–105, 2004.
- [12] 株式会社キャリアタス. 2026 年卒 9 月後半時点の就職意識調査 Vol.01: キャリタス就活 学生モニター 2026 調査結果. 調査レポート, October 2024. 2024 年 10 月発行.
- [13] 株式会社リクルートマネジメントソリューションズ. 2025 年卒・2024 年卒 大学生の就職活動に関する調査. Technical report, リクルートマネジメントソリューションズ, 11 2024. Accessed: 2025-12-30.
- [14] 高橋南海子, タカハシナミコ. 大学生の就職活動に関する実証的研究の動向と課題. *明星大学明星教育センター研究紀要*, No. 8, pp. 1–15, 2018.
- [15] 盛本品子. 時間選好率および現在バイアス性がオンラインゲーム内コンテンツへの課金行動に与える影響. *行動経済学*, Vol. 11, No. 0, pp. 1–13, 2018.
- [16] 田原朱主加, 堀田結孝. 大学生の就職活動不安とソーシャル・サポートとの関連. *帝京大学心理学紀要*, No. 28, pp. 39–56, 2024.

- [17] 内閣官房. 2025（令和 7）年度卒業・修了予定者等の就職・採用活動に関する要請, 4 2024. Accessed: 2025-11-09.
- [18] 内閣府政策統括官（経済財政運営担当）. 学生の就職・採用活動開始時期等に関する調査報告書（令和 6 年度）. Technical report, 内閣府, 12 2024. (参照 2025-11-19).
- [19] 矢崎裕美子, 斎藤和志, 高井次郎. 就職に関する情報探索行動尺度の作成. 名古屋大学大学院教育発達科学研究科紀要, 心理発達科学, Vol. 54, pp. 127–134, 2007.

公開研修参加者が変化を実感する要因とは： 満足度調査からの示唆

永瀬 純[†] 野口 怜[‡]

[†] ‡ 東京理科大学経営学部 〒102-0071 東京都千代田区富士見 1-11-2

E-mail: [†] 8722140@ed.tus.ac.jp, [‡] rnoguchi@rs.tus.ac.jp

あらまし 近年、日本において人材を「資本」と捉え企業価値の向上に繋げていく「人的資本経営」が注目されている。その中で研修を含む人材育成分野は、近年のリスキリング・学び直しの潮流を背景に、重要性が一層高まっている。一方で、現代の研修はスキル向上やキャリア自立を促すことを目的としているものの、その効果が業績に直結しにくく形骸化するリスクも抱える。受講が義務的な消化に終わり、表面的な理解にとどまって主体的な学びや実践に活かされないことが課題として指摘されている。また、人材データは公開されにくく、研修効果に関する実証研究は発展途上である。そこで本稿では、伝統的な研修効果測定手法のカークパトリックの4段階評価モデルを一部改良し、大規模な研修満足度調査データを用いて研修効果が高い受講者の特徴を明らかにした。分析の結果、受講者が自主性に基づいて参加の意思決定をしていること、受講に際して金銭的な負担を伴っていることの2点が重要であることが示唆された。今後はこれら2つの要因についてより詳細な分析を行うことで、研修効果の向上に資する知見が得られると期待される。

キーワード 研修効果測定, 人材育成, アンケート分析

1. はじめに

近年、日本において人材を「資本」ととらえ企業価値の向上に繋げていく「人的資本経営」の重要性が高まっている。企業の原動力は人であり、企業価値の多くは人材の能力やスキル、エンゲージメントによって占められていると言える。実際、人的資本の開示や投資が注目され始めており、2020年には経済産業省による人材戦略に関する経営指針「人材版伊藤レポート」が発表され、人材を資本ととらえ戦略的に育成・活用していくことの重要性が広く認知された[1]。世界経済フォーラムによると、人的資本経営はCapacity（能力ストック）、Development（人材育成・能力開発）、Deployment（配置・活用）、Know-how（知識・技能の応用）の枠組みで考えられる[2]。特に人材育成は、近年のリスキリング・学びなおしの流れの中で、非常に重要視されている。実際、矢野経済研究所の「企業向け研修サービス市場に関する調査」では、人的資本経営に取り組む企業が増え、教育投資の拡充が市場拡大を後押ししていると指摘されている[3]。従業員の能力向上や次世代人材の育成、組織力強化に向けた企業の教育投資は今後も継続すると見込まれ、研修の重要性は今後さらに高まると考えられる。一方で、現代の研修はスキル向上やキャリア自立を促しているものの、効果が業績に直結しにくく形骸化するリスクも抱える。受講が義務的な消化に終わり、表面的な理解にとどまって主体的な学びや実践に活かされないことが課題という指摘もある。実際、知識習得までは評価されていても行動変化まで評価していない研修が多く、研修を

行っても知識が増えただけで終わり、仕事場で活かされて業績に至るまで到達していないケースが多いという報告がある[4][5]。

このように、研修は人的資本投資として大きな期待を背負いながらも、効果が十分に発揮されないリスクを抱えている。したがって研修の効果をより高めるためには、単なる満足度の把握にとどまらず、その基盤となる学びに着目し、研修の効果を具体化するとともに、その影響因子を明らかにすることが重要である。

そこで本研究では、研修効果を「理解」「気持ちの変化」「業務効率の変化」という三つのレベルに整理して測定・分析することで、研修における学びのプロセスと成果との関係を明らかにし、研修設計や運営の改善に資する示唆を得ることを目的とする。

2. 関連研究

2.1 研修効果測定の評価モデル

従来の研修効果研究では、カークパトリックの四段階評価モデルが広く用いられている[6]。本モデルは研修の効果を「反応」「理解」「態度・行動の変容」「組織への影響」という4つの段階から多面的に捉える枠組みであり、実務への活用のしやすさから長年にわたって広く用いられてきた。実際に、学術研究のみならず企業実務においても幅広く採用されており、研究結果を実践に接続しやすいという利点を有する。さらに、各段階が具体的な指標と対応づけやすく、アンケート調査や行動データなど多様なデータを用いた実証分析に適している点も特徴である。

2. 2 研修効果の測定例

先行研究では、集合研修で学んだ内容を職場で実践できるようになる要因を明らかにする調査が行われた[7]。同研究では、効果的な研修運営を目的として、カークパトリックの「態度・行動の変化」に着目して上司が行うフォローアップの重要性を検証する分析を行っている。その結果、フォローアップは常に様に効果をもたらすわけではないことや、研修の目的を明確化することが重要であること、などの示唆を得ている。一方で、同研究では「態度・行動の変化」に対しては主に行動の変化に焦点を当てており、受講者の意識や気持ちの変化については十分に検討されていない。

3. 分析手法

3. 1 使用データ

3. 1. 1 データの収集

本研究で使用するデータは、国立情報学研究所のIDR データセット提供サービスにより株式会社 oricon ME から提供を受けた「企業研修：公開プログラム」の満足度調査データである[8]。オリコンは各種顧客満足度調査を実施しており、本研究ではそのうち研修に関する調査結果を分析対象とした。データは全国の研修受講者を対象にアンケート形式で収集されたものであり、受講内容、満足度、効果に関する項目が含まれている。なお、本研究では提供を受けた匿名化済みデータを二次分析する形で用いた。

3. 1. 2 データの概要

取得したデータから扱う項目を抽出した。12 カテゴリ 94 の変数を使用した。カテゴリは、属性を表すもの（性別・年齢・回答者とその家族の業種・回答者の職業・世帯年収・個人年収・結婚の有無・同居している人）と様々な項目ごとの満足度（満足度評価）、研修のモチベーションに関係があると考えられるもの（参加のきっかけ・費用の負担方法）がある。満足度評価に関しては、20 の総合満足度に加え、それを構成する 14 項目の細かい満足度がある。これらは 4 つのグループに分けることができ、それらは、「学び」、「講師の質」、「研修の運営」、「教材」である。今回の論文では、研修の効果を高める示唆を得るという目的に沿って、学びに関して注目し分析を行う。表のとおり研修効果のレベルを、「理解→気持ちの変化→業務効率の変化」と分けたが、これを先ほどの満足度における学びグループとつなげると「研修内容が参考になったか」がレベル 1、「研修後の気持ちの変化」がレベル 2、「研修後の業務効率の変化」がレベル 3 となる。

3. 2 研修効果のレベル定義

研修効果測定の「気持ちの変化」と「業務効率の変化」の要素をそれぞれ定量的に表すために、カークパトリックの 4 段階評価モデルに基づき、データセット中の満足度小項目である「研修後の気持ちの変化」と「研修後の業務の変化」を代替変数として扱った。カークパトリックの 4 段階評価モデルと本研究で用いる研修効果レベルとの対応関係を以下の表 1 に示す。

表 1 カークパトリックモデルと本研究で用いる研修効果レベルの比較

カークパトリックモデル		本研究	
第 1 段階	研修に対する反応 研修に満足したか	(本研究では研修効果を直接表さない総合満足度は分析対象外)	
第 2 段階	理解度 研修内容を理解できたかどうか	レベル 1	理解度 同様に知識の理解度を確認
第 3 段階	態度・行動の変化 学んだことを実務で実際に使ったか	レベル 2	気持ちの変化 「認識的变化」から「感情的変化」へ焦点化
		レベル 3	業務効率の変化 実務で効率が上がったか (行動そのものではなく、その成果としての効率改善に注目)
第 4 段階	組織への影響 組織全体の業績や成果に与える効果	(本研究では外部要因に左右されやすい組織レベルは扱わず、個人レベルに焦点化)	

3. 3 分析手順

分析手順は以下の 3 段階で行った。
第一に、受講者の変化に影響を与える因子の探索を行った。
第二に、重要因子の関係について確認した。
第三に、研修効果が高いと考えられる理想的な受講者の属性について特定した。以下に、各手順について記述する。

3. 3. 1 受講者の変化に影響を与える因子の探索
本研究では、目的変数を満足度という 1 点から 10 点のカテゴリ変数にするため、3 段階以上の順序付きカテゴリを扱う順序ロジスティック回帰を用いた。これによりどのような要素がそれぞれの学びレベルに影響を与えるのか、特徴量の全体的・平均的な効果を明らかにした。目的変数は、気持ちの変化、と業務効率の変化であり 2 つのモデルを構築した。それぞれ用いた説明変数は同じであり、エリアや性別、回答者とその家族の業種、回答者の職業、結婚の有無などといった属性カテゴリや、参加のきっかけ、費用の負担方法などの研修参加形態に関するカテゴリを用いた。

3. 3. 2 重要因子の関係についての確認

順序ロジスティック回帰の結果より抽出した重要だと思われる因子について、回帰分析ではとらえられ

にくいカテゴリ間の関係性や傾向を補足的に検討するため、クロス集計で分布を確認した。

3. 3. 3 理想的な受講者属性の特定

クロス集計により重要性が示唆された因子について、どのような属性を持つ受講者が該当するかを明らかにするため、ロジスティック回帰を行った。これにより、研修効果が高い受講者の属性的特徴をより明確に把握することを試みた。

4. 分析

4. 1 受講者の変化に影響を与える因子の探索

順序ロジスティック回帰の結果の上位一部を表2に示す。

表 2 順序ロジスティック回帰における影響度上位の変数

説明変数\目的変数	係数		P値		オッズ比	
	気持ち	業務効率	気持ち	業務効率	気持ち	業務効率
自主的に参加	1.4603	1.1709	0	0	4.3074	3.2247
同僚に誘われて参加	0.8176	0.6144	0	0	2.2651	1.8485
父母同居	0.2894	0.3317	0.002	0	1.3356	1.3933
上司の命令で参加	0.4489		0.07		1.5666	
一部自己負担		0.4553		0.01		1.5765
会社員	-0.4591	-0.3909	0.01	0.004	0.6318	0.6764
公務員	-0.4468	-0.4414	0.011	0.012	0.6396	0.6431
甲信越・北陸	-0.3587	-0.3386	0.011	0.017	0.6905	0.7127
東海	-0.3487	-0.2750	0.002	0.013	0.7056	0.7595
関東	-0.1885		0.033		0.8281	
金融		-0.2225		0.006		0.8005

「気持ちの変化」と「業務効率の変化」のいずれにおいても、研修参加における自主性の度合いが高い受講者ほど研修効果が高い傾向が確認された。また、受講料の負担方法や職種・業種が研修効果に関係する可能性が示唆された。

特に「自主的に参加」した受講者が変化を起こしやすいと示唆されたため、以下では研修参加のきっかけに着目した分析を行った。加えて、受講料の負担については、「全額自己負担」や「全額会社負担」よりも「一部自己負担」の影響が相対的に大きいことに関して、事前の想定とは異なる結果が得られた。

この点についても、その要因を検討するため、参加のきっかけと同様に追加的な分析を行った。

4. 2 重要因子の関係についての確認

「気持ちの変化」と「業務効率の変化」について、順序ロジスティック回帰において重要であると示唆された「参加のきっかけ」と「受講料の負担方法」で分けて分布を確認するためのクロス集計の結果を表3と表4に示す。

表 3 「気持ちの変化」における参加のきっかけ×費用負担方法

負担方法 ／参加きっかけ	自主的に参加	同僚に誘われて参加	上司の命令で参加
全額自己負担	7. 4888	7. 2592	6. 2592
全額自己負担（法人割引）	7. 3333	6. 4210	6. 8571
会社＋一部自己負担	7. 7809	7. 1173	6. 8518
全額会社負担	7. 5335	7. 0052	6. 5264

表 4 「業務効率の変化」における参加のきっかけ×費用負担方法

負担方法／ 参加きっかけ	自主的に参加	同僚に誘われて参加	上司の命令で参加
全額自己負担	7. 4333	7. 3703	6. 2592
全額自己負担（法人割引）	7. 0606	6. 6315	7. 0714
会社＋一部自己負担	7. 7243	7. 1173	6. 6543
全額会社負担	7. 4307	6. 9876	6. 4056

「自主的に参加」「同僚に誘われて参加」「命令参加」の順に大別され、その中でも自主的に参加では一部自己負担での受講が最も高かった。また、一部自己負担は他の参加きっかけでも高水準であり、全額会社負担が最上位となるケースは見られなかった。

これらの結果から、自主性の度合いが高いほど受講後の変化が生じやすい傾向であることが分かった。この知見を踏まえ、自主的に研修に参加する受講者がどのような属性的特徴を持つのかについて、データセットから得られる範囲で分析を行うこととした。

4. 3 理想的な受講者属性の特定

変化が生じやすいと示唆された自主性をもって研修に参加する人の特徴を明らかにするため、「自主的に参加」を目的変数にしたロジスティック回帰分析を行った。その推定結果の一部を表5に示す。

表 5 「自主的に参加」を目的変数にした
ロジスティック回帰の影響変数

	変数名	係数	P 値	オッズ比
1	全額自己負担	1. 7154	0. 000	5. 559127
2	会社一部負担	1. 2773	0. 000	3. 586899
3	法人価格	1. 2166	0. 000	3. 375601
4	子供と同居	0. 3396	0. 000	1. 404337
5	金融業	-0. 2982	0. 006	0. 742139

自主的に参加には全額会社負担以外の負担方法（全額自己負担・一部自己負担・法人価格）は「自主的に参加」に対して、正の影響を示した。対照分析として「命令で参加」を目的変数とし、同一の説明変数を用いてロジスティック回帰を行った結果、職種や業種、エリアといった属性要因が影響している一方で、受講料の負担方法が影響している可能性が低いことが示された。

これらの結果から、研修参加における自主性の度合いには、受講料の負担が関係している可能性が示唆された。

5. 考察

5. 1 自律的な学習動機的重要性

これらの結果は、従業員の「自律的な学習動機」をいかに高めるかが、研修効果の向上において重要な鍵となることを示唆している。Deci and Ryan の自己決定理論によれば、人は自らの意思で行動を選択していると認識できる状況において、より高い学習意欲と成果を示すとされている[9]。本研究において「自主的に参加」した受講者ほど研修後の変化が大きかったという結果は、この理論的枠組みと整合的である。

また、受講料の負担方法に関して、全額会社負担よりも自己負担（全額・一部・法人価格）を伴う形態において研修効果が高い傾向が確認された点は、金銭的な投資が研修参加における「自己選択の証拠」として機能し、自律的な学習動機を補強する可能性を示している。すなわち、一定の自己負担を伴うことで、受講者自身が研修参加を能動的に選択したという認識が強まり、学習への関与度が高まると考えられる。この点について、Bietenbeck らは、自己負担比率の違いによって効果の大きさは変動するものの、金銭的投資を行うことで投資回収意識が高まり、研修をより真剣に受講するようになる結果、学習効果が向上する可能性を指摘している[10]。一方で、サンクコスト回収の観点からは、自己負担額が大きいほど研修効果が一樣に高まることも想定されるが、本研究の結果は必ずしもそのような単調な関係を示していない。このことは、金銭的負担が一定水準を超える場合には、心理的負担や期待水準の上昇といった要因が作用し、学習効果との関係が非線形となる可能性を示唆している。

本研究の結果は、こうした先行研究の知見とも概ね整合的であり、研修制度設計において受講料の負担方法が、受講者の動機づけを媒介として研修効果に影響を及ぼし得ることを示唆している。

5. 2 研修効果の向上に資する研修設計の施策案

これらの示唆から、実務的に活用可能な知見は2点得られた。

1 点目は「自主的な参加を促すこと」の重要性である。本研究では、自主的に研修へ参加した受講者ほど研修後の変化が大きいたことが示されており、この結果は、自己決定理論における「自律的な選択が高い学習効果を生む」という知見と整合的である。自主性は研修効果に直接的に寄与する一方で、組織側がどのように介入し、受講者の自主的判断を引き出すかという制度設計上の課題も併せて示している。

2 点目は「受講料を会社が全額負担しないこと」である。受講者に一定の自己負担を求める研修制度は、向上心の高い人材を選別する機能を持ち、結果として

研修効果に間接的に寄与する可能性が示唆された。また、受講料の自己負担は研修参加における「自己選択の証拠」として機能し、受講者の自律的な学習動機を補強し、学習への関与度や真剣度を高める可能性がある。

以上より、研修効果の向上には、従業員の「自律的な学習動機」が重要な役割を果たしていることが示唆される。実務的には、「自主性を引き出す仕組み」と「一定の自己負担を伴う研修費用設計」を両立させた制度設計が有効であると考えられる。具体的には、参加者が研修を自ら選択できる余地を残しつつ、過度ではない自己負担を伴う研修制度を設計することが、研修効果を高める一つの方策となり得る。

5. 3 今後の展望

今後の展望としては、自主的な参加の促進にどのような介入が有効かを検証する必要がある。たとえば、上司からの推薦やキャリア面談での支援が自主性を高めるのか、それとも自律的な選択の妨げとなるのかといった点について、さらなる定量的・定性的研究が求められる。また、どの程度の受講料負担が最も効果があるのかといった点や、研修効果と自主性・費用負担方法の因果関係についてもさらなる分析の余地がある。

本研究では提供していただいたデータを用いて研修効果に関する分析を行ったが、大学の講義などの学生の学習データを収集・分析することも、学習・教育の観点から有益な示唆が得られると考える。具体的には、履修のきっかけや授業参加率、取り組んだ課題の成果などを用いて、自主性と受講前後の変化との関係を明らかにできる可能性がある。アンケート設計をおこなうことで、アンケートデータを用いた分析の発展にも寄与できると考えられる。

また、本研究においてはテキストデータを用いた補足的な分析も行った。しかし、自由記述回答には短い文章が多く、文脈をもった意味のある記述が限られていたため、有益な示唆を得ることが難しい側面があった。さらに、大量の自由記述をその他回答の情報と照合しながら分析するには相応の時間を要するという課題も確認された。これらの課題が克服され、大量の短文自由記述を迅速かつ体系的に処理することが可能となれば、従来の量的分析では捉えきれなかった視点からの仮説構築や、新たな示唆の創出につながる可能性がある。

6. おわりに

本研究は、効果的な研修のあり方を明らかにすることを目的に、研修の満足度を分析したものである。分析で得られた知見として、「受講者の自主性」と「受講

者の金銭負担」が重要であると示唆された。これより、従来そのように想定されてきた事柄が、データ分析を通じてより明確になったと考えられる。加えて、本研究の結果は、現場での実践や今後の更なる分析の基礎となることが期待される。また、本研究では主に量的分析を用いたが、今後はアンケートにおける短文自由記述データの活用が挙げられる。特に、1文程度の短い自由記述を大量に収集する形式のアンケートは、回答者の負担が小さいためデータを集めやすく、実務上も有用であると考えられる。1文アンケートのような大量の短文コメントを有益な情報を保持したまま、迅速に分析することが可能となれば、人事領域にとどまらず、マーケティングなど他分野への応用可能性も広がると考えられる。

7. 謝辞

本研究の中核となる貴重なデータを提供いただいた株式会社 oricon ME および、国立情報学研究所の IDR データ提供サービスに心より感謝申し上げます。

参 考 文 献

- [1] 伊藤邦雄, ”人材版伊藤レポート”, 2020.
- [2] 世界経済フォーラム, ”Measuring Human Capital”, 2021.
- [3] 矢野経済研究所, ”企業向け研修サービス市場に関する調査”, 2025.
- [4] Dita Nīmantel, Maija Kokare, Sanita Baranova1, and Svetlana Surikova, ”Transferring Results of Professional Development into Practice: A Scoping Review”, 2025.
- [5] Shishi Wu, Imara Roychowdhury and Mishal Khan, ”Evaluations of training programs to improve human resource capacity for HIV, malaria, and TB control: a systematic scoping review of methods applied and outcomes assessed”, 2017.
- [6] Kirkpatrick, D. L. ”Techniques for evaluating training programs”, Journal of the American Society of Training Directors, 13, 21-26, 1959.
- [7] 小 菌 修, ”効果的な社員研修についての考察”, 2008.
- [8] 株式会社 oricon ME ”顧客満足度調査データ”, 国立情報学研究所情報学研究データリポジトリ, 2024. <https://doi.org/10.32130/idr.10.1>.
- [9] Edward L. Deci and Richard M. Ryan, ”Intrinsic motivation and self-determination”, 1985.
- [10] Jan Bietenbeck, Andreas Leibing, Jan Marcus and Felix Weinhardt, ”Tuition fees and educational attainment”, 2023.

段階評価データの分布形状差に着目した非類似度指標と ウェルビーイングデータ分析への応用

萩田 碧偉[†] 山崎綾一郎[†] 山岸 祐己^{†,††,†††} 青木 成樹^{††††} 橋本 正洋^{††††,†††††,††††††}

[†] 静岡理工科大学 〒420-0857 静岡県静岡市葵区御幸町 20 M20 11F

^{††} 浜松医科大学 〒431-3129 静岡県浜松市中央区半田山 1-20-1

^{†††} 株式会社良品計画 〒112-0004 東京都文京区後楽 2-5-1

^{††††} マリンオープンイノベーション機構 〒424-0922 静岡市清水区日の出町 9 番 25 号 清水マリビル 2 階

^{†††††} 法政大学 〒102-8160 東京都千代田区富士見 2-17-1

^{††††††} 東京科学大学 〒152-8550 東京都目黒区大岡山 2-12-1

E-mail: †{2218089.ha,2421028.yr,yamagishi.yuki}@sist.ac.jp

あらまし 本研究では、段階評価(リッカート尺度)データにおける属性集団間の評価傾向の差を、平均などの単一統計量ではなく評価分布の「形状差(偏り・裾の重さ)」として捉えるため、Jarque-Bera 統計量に基づく歪度・尖度差を標準化して統合した非類似度指標 Jarque-Bera 距離を提案する。さらに、属性別集計で避けがたい小標本・不均衡標本の影響を扱うため、有効標本サイズに基づく標本サイズ補正(D_{JB-SD})と、上限制御を導入した安定化版($D_{JB-SDcap}$)を併せて定式化する。提案指標について、まずシミュレーションにより偽陽性率と検出力を評価し、「形状差をどの程度安定に検出できるか」と「標本規模差が挙動にどう現れるか」を整理する。次に MovieLens 評価データを用いて、距離行列の整合性とクラスタ構造の整合性を比較し、均衡条件では比較手法と整合しつつ、不均衡条件では補正の設計に由来する差分が距離構造として顕在化することを示す。最後に静岡市ウェルビーイングデータへ適用し、歪度・尖度に基づく解釈可能な属性クラスタが得られること、ならびに標本サイズ補正が小標本属性の過敏な孤立を抑え得ることを確認し、行政アンケートに対する実務的な適用可能性を検証する。

キーワード 非類似度指標, 分布間距離, モーメント, リッカート尺度, ウェルビーイング

1 はじめに

行政や企業などにおいて、住民・顧客の満足度や意識を把握するためのアンケート調査は広く用いられている。これらの調査では、回答負担の低さと集計の容易さから、1〜5 などの段階からなる段階評価(リッカート尺度[5])が採用されることが多い。

一方で段階評価データは、属性集団ごとや設問項目ごとに観測される評価分布の形状が異なり、平均値等の単一の基本統計量だけでは、偏りや尖り、二極化といった分布形状の差異を十分に捉えられない課題がある。したがって、段階評価アンケートの分析においては、単一の基本統計量に依存せず、分布の形状差を直接比較できる枠組みが求められている。

多数の分布を横断的に比較する枠組みとして、非類似度指標が用いられる。これらは分布全体の差異を単一尺度に要約でき、属性集団や設問項目を横断した比較・可視化に有用である。一方で既存指標の多くは、算出手順が複雑であるため、非専門家においては解釈性や説明性に懸念が残る。また、パラメータ設定の必要性に加え、属性別集計では各集団の標本サイズが小さくなりやすく、さらに集団間で標本規模が不均衡になりやすい。このとき、各集団で分布の推定精度(ばらつき)が異なるため、偶然に生じた差を差異として過大評価した比較につながるおそ

れがあり、標本サイズの影響を明示的に扱える枠組みが求められる。

本研究では、歪度・尖度という少数の形状指標に基づき、Jarque-Bera(JB) 統計量[3]の考え方を「2 標本の分布差」へ拡張した Jarque-Bera (JB) 距離を提案する。提案手法は特に行政などの実務利用を想定し、(i) 結果の解釈性・説明性が高いこと、(ii) 小標本・不均衡標本における推定誤差(標本サイズの影響)を距離設計として明示的に取り込めることを狙いとする。

本稿ではまず、提案する JB 距離と比較手法(CMD, W_1)を述べる。次に、シミュレーションにより段階評価データでの偽陽性率・検出力を確認する。続いて、MovieLens データで比較手法との距離構造・クラスタ構造の整合性を評価し、実データにおける妥当性と差分を整理する。最後に、静岡市ウェルビーイング調査へ適用し、実務データでの解釈可能性と運用上の示唆を述べる。

2 提案手法

2.1 JB 統計量

Jarque-Bera(JB) 統計量は正規性検定のための統計量であるが、その構造は「分布形状を表す歪度・尖度を、推定誤差(標準誤差)で標準化した上で統合する」設計として理解できる。JB 統計量は

$$JB = \frac{n}{6}S^2 + \frac{n}{24}(K-3)^2 \quad (1)$$

で与えられる [3]. ここで S は標本歪度, K は標本尖度 (正規分布の基準値は $S = 0, K = 3$) である.

正規分布下では $\text{Var}(S) \approx 6/n, \text{Var}(K) \approx 24/n$ が成り立つため, 標準誤差は $SE(S) \approx \sqrt{6/n}, SE(K) \approx \sqrt{24/n}$ と近似できる.

このとき, 正規分布の基準値からのずれを標準誤差で割った標準化量 (Z スコア) として

$$\begin{aligned} Z_S &= \frac{S-0}{\sqrt{6/n}} = \sqrt{\frac{n}{6}} S, \\ Z_K &= \frac{K-3}{\sqrt{24/n}} = \sqrt{\frac{n}{24}} (K-3) \end{aligned} \quad (2)$$

を導入すると, 式 (1) は

$$JB = Z_S^2 + Z_K^2 \quad (3)$$

と書ける. すなわち JB 統計量は, 「歪度・尖度のずれを推定誤差で標準化し, それらの大きさを二乗和として統合した量」である. 本研究ではこの考え方を, 正規分布からの乖離度の定量化ではなく, 2 つの分布間の形状差の定量化へ拡張する.

2.2 2 標本分布への拡張: 形状差の標準化

2 つの標本分布 X, Y の標本サイズを n_X, n_Y , 標本歪度を S_X, S_Y , 標本尖度を K_X, K_Y とする. 分布形状の差は歪度差 ($S_X - S_Y$) および尖度差 ($K_X - K_Y$) で捉えられるが, 小標本では S や K の推定誤差が大きく, 単純な差のみでは偶然差を過大評価し得る. そこで JB 統計量と同様に, 観測された形状差を推定誤差で標準化して評価する.

2 標本が独立であると仮定すると, 分散の加法性より差の分散は

$$\begin{aligned} \text{Var}(S_X - S_Y) &\approx \frac{6}{n_X} + \frac{6}{n_Y}, \\ \text{Var}(K_X - K_Y) &\approx \frac{24}{n_X} + \frac{24}{n_Y} \end{aligned} \quad (4)$$

と近似できる. したがって, 2 標本の歪度差・尖度差に対する標準化量を

$$\begin{aligned} Z_{(S_X - S_Y)} &= \frac{S_X - S_Y}{\sqrt{6/n_X + 6/n_Y}}, \\ Z_{(K_X - K_Y)} &= \frac{K_X - K_Y}{\sqrt{24/n_X + 24/n_Y}} \end{aligned} \quad (5)$$

と定義できる. これらは「観測された形状差が, 推定誤差 (標準誤差) に対してどの程度大きい」を表す量であり, 小標本・不均衡標本では分母が大きくなるため, 同じ形状差でも過度に大きく評価されにくいという直観と整合する.

2.3 提案指標: 標本サイズ補正付非類似度指標

前節で導入した標準化差 $Z_{(S_X - S_Y)}$ および $Z_{(K_X - K_Y)}$ に基づき, 本研究では 2 標本分布間の形状差を表す非類似度指標を提案する. 本稿ではこの指標ファミリーを **JB 距離 (Jarque–Bera dissimilarity measure ; D_{JB})** と総称し, まず標本

サイズの影響を考慮する主指標として D_{JB-SD} (**sample-size dependent ; 標本サイズ依存型**) を定義する.

本手法では, 標準化差ベクトル $(Z_{(S_X - S_Y)}, Z_{(K_X - K_Y)})$ を単一尺度に要約するため, その L_2 ノルムとして

$$D_{JB-SD}(X, Y) = \sqrt{Z_{(S_X - S_Y)}^2 + Z_{(K_X - K_Y)}^2} \quad (6)$$

を定義する.

式 (6) を展開すると, 分母に $\frac{1}{n_X} + \frac{1}{n_Y}$ が現れることが分かる. ここで調和平均に基づく有効標本サイズ

$$n_{\text{eff}} = \frac{2}{\frac{1}{n_X} + \frac{1}{n_Y}} \quad (7)$$

を導入すると, 式 (6) は

$$D_{JB-SD}(X, Y) = \sqrt{\frac{n_{\text{eff}}}{12} (S_X - S_Y)^2 + \frac{n_{\text{eff}}}{48} (K_X - K_Y)^2} \quad (8)$$

と整理できる.

式 (8) は, 歪度差・尖度差を推定誤差に基づくスケーリングで調整しつつ, それらを単一の非負尺度として統合していることを明確にする. 特に n_{eff} は調和平均に基づくため, 不均衡標本では小さい方の標本サイズの影響が大きく反映され, 小標本側の不確かさが距離に与える過大評価を抑制しやすい. なお, 複数分布を横断的に比較する場合, n_{eff} は分布対ごとに異なり得るため, 不均衡標本を含む比較ではスケーリング差が距離構造に影響し得る. 一方で, 標本規模が概ね同程度の比較ではスケーリングはほぼ共通となり, 形状差そのものと同様の距離構造を与える場合がある.

また, 本指標は非負性, 対称性 $D_{JB-SD}(X, Y) = D_{JB-SD}(Y, X)$ を満たす一方で, 同一性および三角不等式を常に満たす保証はないため, 厳密には距離ではなく非類似度指標として位置づけられる. ただし, 文脈に応じて広義の「距離」と表現する場合がある.

2.4 提案指標: 参照指標と上限制御付き指標

形状差そのものを標本サイズに依らず同一基準で比較するため, 参照指標として D_{JB-SI} (**sample-size invariant ; 標本サイズ不変型**) も併用する. D_{JB-SI} は式 (8) においてスケーリングを除いた形であり,

$$D_{JB-SI}(X, Y) = \sqrt{\frac{1}{12} (S_X - S_Y)^2 + \frac{1}{48} (K_X - K_Y)^2} \quad (9)$$

と定義する. 本指標は, 形状差 (歪度差・尖度差) のみを評価する基準として用いる.

さらに, 大標本領域では推定誤差の改善が逓減し得るにもかかわらず, D_{JB-SD} が n_{eff} により機械的に増大し続ける可能性がある. そこで, 本研究ではスケーリングを上限で打ち切る $D_{JB-SDcap}$ (**sample-size dependent with capped ; 上限制御付き標本サイズ依存型**) を導入する. $\tilde{n}_{\text{eff}} = \min(n_{\text{eff}}, n_{\text{cap}})$ として

$$D_{JB-SDcap}(X, Y) = \sqrt{\frac{\tilde{n}_{eff}}{12}(S_X - S_Y)^2 + \frac{\tilde{n}_{eff}}{48}(K_X - K_Y)^2} \quad (10)$$

を導入する。

2.5 比較手法：CMD および W_1

本稿では、提案手法の比較対象として、Central Moment Discrepancy(CMD) [8] と Wasserstein-1 距離 (W_1) [7] を用いる。いずれも標本分布 (経験分布) から算出可能な分布間距離 (非類似度指標) である。

■ CMD

CMD は、平均および高次中心モーメント差を次数ごとに評価して統合する指標である。標本 $X = \{x_i\}_{i=1}^{n_X}$, $Y = \{y_j\}_{j=1}^{n_Y}$ に対し、標本平均 $\bar{x}, \bar{y}$ と deg 次中心モーメント

$$C_{deg}(X) = \frac{1}{n_X} \sum_{i=1}^{n_X} (x_i - \bar{x})^{deg}, \quad C_{deg}(Y) = \frac{1}{n_Y} \sum_{j=1}^{n_Y} (y_j - \bar{y})^{deg} \quad (11)$$

を用いると、有限区間 $[a, b]$ で正規化した CMD は

$$\begin{aligned} \text{CMD}_{\text{Deg}}(X, Y) = & \frac{1}{|b - a|} \|\bar{x} - \bar{y}\|_2 \\ & + \sum_{deg=2}^{\text{Deg}} \frac{1}{|b - a|^{deg}} \|C_{deg}(X) - C_{deg}(Y)\|_2 \end{aligned} \quad (12)$$

として計算できる。本稿では既定値に従い $Deg = 5$ を用い、区間 $[a, b]$ は実験設定に応じて与える。

■ W_1

W_1 は最適輸送 [4] に基づく分布間距離であり、本稿では経験分布

$$\hat{P}_X = \frac{1}{n_X} \sum_{i=1}^{n_X} \delta_{x_i}, \quad \hat{P}_Y = \frac{1}{n_Y} \sum_{j=1}^{n_Y} \delta_{y_j} \quad (13)$$

に対する $W_1(\hat{P}_X, \hat{P}_Y)$ を用いる。コスト関数は $C(x, y) = |x - y|$ とする。

1次元では最適輸送に単調性が成立するため、線形計画を解かずとも厳密に計算できる。具体的には、 $X \cup Y$ の昇順ユニーク値を $z_1 < \dots < z_L$ とし、各点の質量を $p_\ell = \frac{1}{n_X} \#\{i : x_i = z_\ell\}$, $q_\ell = \frac{1}{n_Y} \#\{j : y_j = z_\ell\}$ とおく。累積質量 $P_\ell = \sum_{u=1}^{\ell} p_u$, $Q_\ell = \sum_{u=1}^{\ell} q_u$ を用いると

$$W_1(\hat{P}_X, \hat{P}_Y) = \sum_{\ell=1}^{L-1} |P_\ell - Q_\ell| (z_{\ell+1} - z_\ell) \quad (14)$$

と表される。式 (14) は、標本を昇順にソートし、小さい方から質量を対応付けてコストを積算する (ソート + 貪欲マッチング) 計算と等価である。

3 評価実験 1：シミュレーションによる偽陽性率・検出力評価

本節では、提案する JB 距離 (D_{JB-SD}, D_{JB-SI}) が、分布形

状差 (歪度・尖度) をどの程度検出できるかを、シミュレーションにより検証する。特に、(i) 分布差が存在しない帰無条件における誤検出の生じやすさ (偽陽性率), および (ii) 形状差が存在する対立条件における検出のしやすさ (検出力) を評価する。比較対象として CMD(区間 [1,5]) および W_1 を用いる。

3.1 設定

段階評価データを想定し、値域 [1, 5] の離散値標本 X, Y を用いる。基準として標準正規分布 (平均 0・標準偏差 1・歪度 0・尖度 3) の分布を取り、帰無条件では同一形状、対立条件では表 2 に従い、標本 Y の歪度・尖度を変化させて連続値標本を生成する。生成した標本は 5 段階に離散化して 1 ~ 5 のデータとする。

サイズ条件は表 1 の 6 条件 (均衡 3, 不均衡 3), 形状差条件は表 2 の 6 条件 (歪度差のみ/尖度差のみ/両方差, それぞれ小・大) とする。なお離散化処理により尖度差が観測上特に小さくなりやすいため、尖度差「大」を相対的に大きく設定した。

表 1 サイズ条件 (n_X, n_Y)

ケース	(n_X, n_Y)
S1	(100, 100)
S2	(1000, 1000)
S3	(10000, 10000)
S4	(100, 500)
S5	(100, 1000)
S6	(100, 10000)

表 2 形状差条件 (小/大)

種別	小	大	備考
歪度差	0.5	1.0	尖度は基準値に固定
尖度差	0.5	2.0	歪度は基準値に固定
両方差	(0.5, 0.5)	(1.0, 2.0)	(歪度差, 尖度差)

3.2 しきい値と評価指標

本節では、距離がしきい値を超えた場合に「分布差あり」と判定し、その超過率として偽陽性率および検出力を推定する。しきい値は帰無条件 (X, Y 同分布) で距離を反復計算し、95 パーセンタイルを採用した。反復回数は各条件で $R = 10000$ 回とする。目的に応じて次の 2 種類のしきい値を用いる：

- **基準しきい値** T_{ref} : 均衡・大標本 (S3) で推定したしきい値を、全サイズ条件 (S1–S6) に共通適用する。これにより、標本規模差 (小標本・不均衡) に対する誤検出の上振れを比較する。
- **条件別しきい値** $T_{(n_X, n_Y)}$: 各サイズ条件ごとに推定したしきい値を、同一条件に適用する。これにより、サイズ条件の違いを揃えたフェアな検出力比較を行う。

偽陽性率は帰無条件下で T_{ref} を超えた割合として推定する。検出力は対立条件下で $T_{(n_X, n_Y)}$ を超えた割合として推定する。なお、 D_{JB-SD} と D_{JB-SI} はスケーリングの関係から $T_{(n_X, n_Y)}$ による評価では等価となるため、本稿ではこれらをまとめて

$D_{JB}(T_{(n_X, n_Y)})$ として扱う。一方で補正の影響を確認するため、 T_{ref} を用いた D_{JB-SD} と D_{JB-SI} の差も併せて評価する。

3.3 結果と考察

■ 偽陽性率

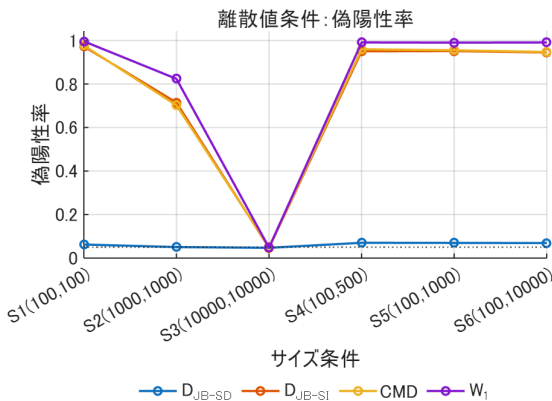


図 1 偽陽性率: 基準しきい値 T_{ref} (S3 で推定), $R = 10000$

図 1 に、偽陽性率の結果を示す。基準しきい値 T_{ref} を全サイズ条件に共通適用した場合、 D_{JB-SD} は S1–S6 の広い範囲で偽陽性率が概ね 0.05 付近に保たれ、標本サイズ補正により偶然差の過大評価が抑制されることが確認できる。一方、 D_{JB-SI} および CMD, W_1 は、基準条件 (S3) ではしきい値設定上 0.05 となるが、小標本・不均衡条件では偽陽性率が大きく上振れし得る。これは、基準条件 (大標本・均衡) で得た比較的小さいしきい値を、推定誤差の大きい条件へ流用するストレステストであるためであり、標本サイズの影響を明示的に扱う必要性を示す結果である。

■ 検出力

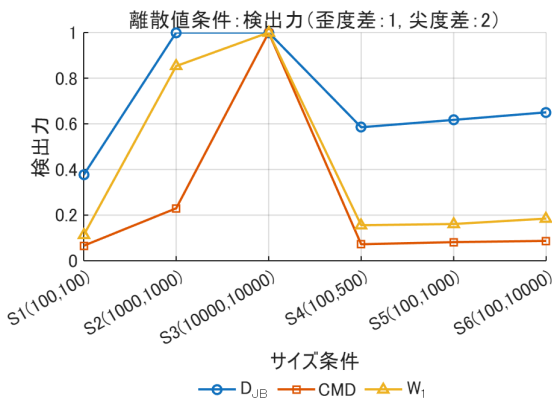


図 2 検出力 (両方差大): 条件別しきい値 $T_{(n_X, n_Y)}$, $R = 10000$

図 2 に、代表的な形状差条件として両方差 (大) の検出力を示す。条件別しきい値 $T_{(n_X, n_Y)}$ を用いたフェアな比較では、 $D_{JB}(T_{(n_X, n_Y)})$ は小標本・不均衡条件を含む多くのサイズ条件で CMD および W_1 と同程度、あるいはそれらを上回る検出力を示した。また、基準しきい値 T_{ref} で D_{JB-SD} と D_{JB-SI} を比較すると、 D_{JB-SI} の方が高い傾向が見られ、標本サイズ補正は偽陽性率を抑制する一方で検出力を一定程度抑えるトレ

ドオフを持つことが分かる。

なお、連続値および形状差条件に加え、平均・分散が変動する連続値混合条件においても同様の傾向を確認しているが、本稿では段階評価データを主対象とするため詳細は省略する。

4 評価実験 2: MovieLens による比較手法との整合性評価

本節では、MovieLens 32M¹ [1] のジャンル別評価分布を用い、提案手法 JB 距離が既存の分布間距離とどの程度整合するかを検証し、整合性 (類似性) および差分 (新規性) を整理することを目的とする。

4.1 データと設定

MovieLens 32M は 0.5–5.0 の 10 段階からなる映画評価データである、ジャンル別に評価分布を構成できる。本実験では、標本サイズ上位 10 ジャンル (G10) を対象とし、各ジャンルから $H = 3$ 個のサブ分布をサブサンプリングで作成する (G10H3)。「ジャンル × サブ分布」(計 30 分布) 間の距離行列を構成し、凝集型階層クラスタリング (群平均法) を適用する。

実験手法は JB 距離 (D_{JB-SD} , D_{JB-SI} , $D_{JB-SDcap}$), CMD (区間 [0.5, 5.0]), および W_1 とする。 $D_{JB-SDcap}$ の上限制御は $n_{cap} = 10000$ (暫定設定) とした。

本実験で用いるサンプルサイズ条件を表 3 に示す。均衡条件 (B1–B3) では各サブ分布の標本サイズを同一とし、不均衡条件 (I1–I3) では各サブ分布の標本サイズを $\{n_1, n_2, n_3\}$ の 3 水準に割り当てて規模差を導入する。

表 3 標本サイズ条件 (G10H3)

種別	条件	標本サイズ (各サブ分布の n)
均衡	B1	$n = 1000$
	B2	$n = 10000$
	B3	$n = 50000$
不均衡	I1	$n \in \{100, 500, 1000\}$
	I2	$n \in \{100, 1000, 10000\}$
	I3	$n \in \{5000, 10000, 50000\}$

以降では、代表として均衡・大標本条件 (B3) と、不均衡条件 (I2, I3) に絞って結果を示す。なお、B3 では JB 距離 3 種は補正係数による定数倍関係となり距離構造・クラスタ構造が一致するため、表では D_{JB-SD} のみを掲載する。また I2 では $n_{cap} = 10000$ の設定により $D_{JB-SD} = D_{JB-SDcap}$ となるため、同様に省略する。

4.2 評価指標

距離構造の整合性は Mantel 相関 r_M [6] により評価する。Mantel 相関は、2 つの距離行列の対応要素 (上三角要素) 間の相関を測る指標であり、値が大きいほど「分布同士の近接関係 (距離構造)」が類似することを意味する。

またクラスタ構造の整合性は、階層クラスタリング結果の一致度を Adjusted Rand Index (ARI) [2] により評価する。ARI

1 : <https://grouplens.org/datasets/movielens/>

は 2 つのクラスタ割当の一致度を偶然一致を補正して測る指標であり、値が大きいほどクラスタ構造が類似することを意味する。本稿では、大域構造に対応する二分割 ($Cl = 2$) に絞って報告する。

さらに補助的に、ジャンル名を外部参照に基づく擬似真ラベルとみなし、復元度 (擬似真ラベル ARI) を評価する。具体的には、同一ジャンル由来のサブ分布同士が優先的に近接し、同一クラスタにまとまりやすいかを測ることで、距離構造が外部ラベル (ジャンル) と整合する度合いを外部妥当性 (補助評価) として確認する。

4.3 結果と考察

■距離構造の整合性 (Mantel 相関)

表 4 距離構造の整合性：Mantel 相関			
手法ペア	B3	I2	I3
(D_{JB-SD}, CMD)	0.974	0.260	0.854
(D_{JB-SD}, W_1)	0.973	0.217	0.860
(D_{JB-SI}, CMD)	–	0.679	0.950
(D_{JB-SI}, W_1)	–	0.613	0.954
$(D_{JB-SDcap}, \text{CMD})$	–	–	0.955
$(D_{JB-SDcap}, W_1)$	–	–	0.958

表 4 に Mantel 相関の結果を示す。均衡条件では、JB 距離と比較手法 (CMD, W_1) の相関が高く、距離構造 (近接関係) が概ね整合する結果を示した (例：B3 では $r_M \approx 0.97$)。この結果は、実データにおいても JB 距離が比較手法と同様の近接関係を一定程度再現し得ることを示し、整合性の観点から妥当性を支持する。一方、不均衡条件 (特に I1, I2) では D_{JB-SD} の整合性が低下するのに対し、 D_{JB-SI} は相対的に高い整合性を維持しており、標本規模差を距離に反映させる設計の違いが距離構造の差として顕在化することが分かる。さらに I3 では、上限制御を導入した $D_{JB-SDcap}$ が比較手法との整合性を高める場合があり (例： $(D_{JB-SD}, \text{CMD}) = 0.854$, $(D_{JB-SDcap}, \text{CMD}) = 0.955$)、大標本域での距離の過度な増大を抑える設計が距離構造の極端化を緩和し得ることが示唆される。

■クラスタ構造の整合性 (手法間 ARI)

表 5 クラスタ構造の整合性：手法間 ARI($Cl = 2$)

手法ペア	B3	I2	I3
(D_{JB-SD}, CMD)	1.000	0.187	1.000
(D_{JB-SD}, W_1)	1.000	0.187	1.000
(D_{JB-SI}, CMD)	–	-0.034	1.000
(D_{JB-SI}, W_1)	–	-0.034	1.000
$(D_{JB-SDcap}, \text{CMD})$	–	–	1.000
$(D_{JB-SDcap}, W_1)$	–	–	1.000

表 5 に手法間 ARI の結果を示す。均衡条件では、大域的な二分 ($Cl = 2$) では一致しやすい一方で、ジャンル粒度 ($Cl = G$) では一致度が中程度に留まる場合がある。(例：B3 で (D_{JB-SD}, CMD) は $Cl = 2$ で 1, $Cl = 10$ で 0.458)。すなわち、均衡条件では大域構造は整合しやすいが、細かな結合順序には手法差が残る。

不均衡条件では、 $Cl = 2$ の一致度が低下し負となる例も見られ (例：G10H3/I2 で $(D_{JB-SI}, \text{CMD}) = -0.034$)、「どのまとまりを先に二分するか」という大域構造から手法差が現れやすい。これは前段の Mantel 相関における距離構造差と整合的であり、不均衡下では距離構造の差分がクラスタ構造の差として可視化されることを示す。

■外部妥当性：復元度 (擬似真ラベル ARI)

表 6 補助評価：擬似真ラベル (ジャンル) に対する復元度 (ARI)

手法	B3	I2	I3
D_{JB-SD}	0.499	0.102	0.328
D_{JB-SI}	–	0.040	0.261
$D_{JB-SDcap}$	–	–	0.252
CMD	0.912	0.102	0.373
W_1	0.865	0.097	0.638

表 6 に復元度 (擬似真ラベル ARI) の結果を示す。不均衡条件では D_{JB-SD} が D_{JB-SI} を上回る (例：I3 で $0.328 > 0.261$) 場面が見られる。距離構造の整合性 (表 4) やクラスタ構造の整合性 (表 5) が D_{JB-SI} より低い条件でも D_{JB-SD} の復元度が高くなり得ることは、 D_{JB-SD} が不均衡標本下で比較手法とは異なる距離構造を形成しつつ、外部参照 (ジャンル) に照らして有利に働き得る差分を持つ可能性を示唆する。一方で $D_{JB-SDcap}$ は整合性を高める一方、復元度が必ずしも向上しない場合もあり (I3 で 0.252)、整合性と外部妥当性の間にトレードオフが生じ得る可能性に留意が必要である。

5 評価実験 3：ウェルビーイングデータへの適用

本実験は、提案手法を行政アンケート (ウェルビーイングデータ) に適用し、実務的な分析道具として有用かを検証する。実務データにおいて、提案手法が (i) 属性集団が「評価傾向の類似性」に基づいて解釈可能なクラスタを形成するか、(ii) 分布形状 (歪度・尖度) に基づく特徴が既存手法 (CMD, W_1) とは異なる形で可視化されるか、(iii) 標本サイズを考慮する D_{JB-SD} 、および上限制御を導入した $D_{JB-SDcap}$ が、小標本属性を含む集計においてどのように振る舞うかを中心に確認する「実務的検証」に位置づけられる。

5.1 データ概要と分析設定 (属性別)

静岡市が令和 6 年 1 月に実施した、16 歳以上の市民を対象とする地域幸福度に関する満足度調査データ (全 47 項目、5 段階) を用いる。評価に欠損のある回答者は除外し、有効回答者数は 2305 名であった。

各回答者は、区別・性別・年代・性年代・職業・家族構成の 6 分類について、それぞれ 1 つずつ属性ラベルを同時に持つ (1 人あたり計 6 属性)。このとき性年代 (例：10 代女性) は性別 × 年代から派生する分類であり、年代別 (10 代) や性別 (女性) と独立ではない。一方で区別・職業・家族構成は性年代ほどの派生関係を持たず、概ね独立な分類として扱える。したがって、属性別分析で扱う各属性群は相互に排反な集合とは限らず、属性群どうしが回答者レベルで重なり得る点に留意する。

属性は全 38 属性に加えて全回答データを全体属性として含めた計 39 種類であり、属性集団の規模は最小 15 人 (10 代男性)、最大 2305 人 (全体) である。

実験手法として D_{JB-SD} , D_{JB-SI} , $D_{JB-SDcap}$, CMD(区間 [1,5]), W_1 を用いる。クラスタリングは距離行列に対して凝集型階層クラスタリング (群平均法) を適用し、クラスタ数は属性の大部分が 6 分類であることを踏まえて実験的に 6 クラスタに固定した (クラスタ数の最適化は今後の課題)。

クラスタの形成理由を解釈するため、回答者数 (N)、標本サイズ (回答者数 $\times$ 項目数)、平均、標準偏差、歪度、尖度、および段階別比率 ($P_1 \sim P_5$) を算出し、併せて参照する。

5.2 結果： D_{JB-SD} による属性クラスタの解釈

図 3 に D_{JB-SD} による属性別デンドログラムを、表??に 6 クラスタに対応する各属性の基本統計量を示す。以下では、デンドログラムで得られたクラスタ構造を起点に、各統計量を用いて「なぜ同一クラスタか」を確認し、実務上のラベル付けを行う。

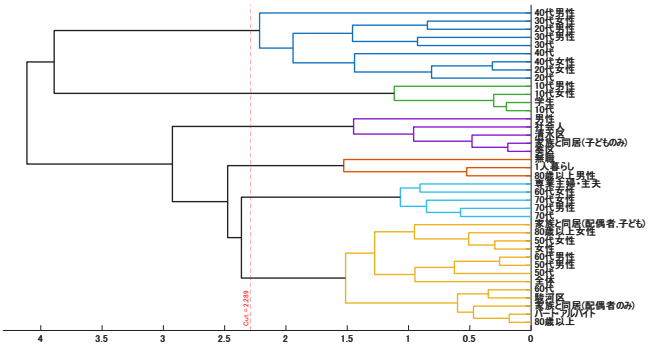


図 3 属性別デンドログラム (D_{JB-SD} , 群平均法, 6 クラスタ)

■ Cl.A1：現役世代 (やや低評価寄り・ばらつき大)

20～40 代と同年代の性年代 (20 代～40 代男女) がまとまる。歪度は相対的に高め (負の偏りが弱い～0 近傍) で、低評価側 (1～2) も一定程度含まれやすい。尖度は相対的に小さめで特定段階への集中が強くない、標準偏差が大きめとなる傾向とも整合する。

■ Cl.A2：若年・学生 (高評価側集中)

10 代, 10 代男女, 学生がまとまる。歪度が強く負になり、高評価側 (4～5) に回答が集まりやすい傾向を示す。本クラスタには $N = 15$ の小標本属性 (10 代男性) が含まれるが、 D_{JB-SD} では小標本ゆえの偶然差が過敏に孤立しにくく、同傾向の若年・学生クラスタに統合されている点が重要である。

■ Cl.A3：代表的属性 (中庸)

葵区・清水区, 男性, 社会人など標本サイズの大きい主要属性がまとまる。歪度・尖度が中程度で分布形状が標準的なまとまりであり、実務的には基準群として他クラスタを相对比较する起点となる。

■ Cl.A4：高齢男性・1 人暮らし・無職 (中心集中・相対的にやや低評価寄り)

80 歳以上男性, 無職, 1 人暮らしがまとまる。歪度は相対的

表 7 D_{JB-SD} による属性別クラスタ (6 クラスタ) と基本統計量 (回答者数, 標本サイズ, 平均, 標準偏差, 歪度, 尖度, 段階別比率 $P_1 \sim P_5$)

クラスタ	属性名	回答者数	標本サイズ	平均	標準偏差	歪度	尖度	P1	P2	P3	P4	P5
Cl.A1	20代	127	5,969	3.160	1.113	-0.245	2.350	0.088	0.183	0.315	0.309	0.105
	30代	230	10,810	3.054	1.135	-0.172	2.257	0.109	0.199	0.314	0.285	0.093
	40代	313	14,711	3.083	1.080	-0.178	2.399	0.087	0.199	0.343	0.287	0.084
	20代男性	43	2,021	3.100	1.180	-0.258	2.228	0.127	0.162	0.305	0.297	0.109
	20代女性	81	3,807	3.213	1.066	-0.220	2.393	0.062	0.195	0.316	0.323	0.104
	30代男性	103	4,841	3.009	1.120	-0.134	2.269	0.110	0.211	0.323	0.274	0.083
	30代女性	127	5,969	3.091	1.145	-0.205	2.254	0.108	0.190	0.307	0.295	0.101
	40代男性	128	6,016	3.042	1.092	-0.112	2.357	0.093	0.212	0.342	0.267	0.086
	40代女性	177	8,319	3.114	1.075	-0.225	2.421	0.084	0.190	0.338	0.303	0.085
Cl.A2	10代	42	1,974	3.444	1.185	-0.444	2.367	0.079	0.133	0.261	0.319	0.208
	10代男性	15	705	3.566	1.210	-0.561	2.415	0.077	0.118	0.230	0.315	0.251
	10代女性	26	1,222	3.396	1.180	-0.427	2.348	0.083	0.142	0.257	0.332	0.186
	学生	59	2,773	3.406	1.185	-0.459	2.367	0.086	0.137	0.250	0.340	0.187
Cl.A3	葵区	849	39,903	3.126	1.047	-0.214	2.499	0.074	0.189	0.356	0.300	0.082
	清水区	746	35,062	3.084	1.052	-0.219	2.520	0.085	0.183	0.370	0.287	0.076
	男性	1,032	48,504	3.098	1.031	-0.205	2.541	0.075	0.190	0.371	0.291	0.073
	社会人	1,002	47,094	3.102	1.057	-0.226	2.476	0.082	0.188	0.353	0.299	0.078
	家族と同居 (子どものみ)	189	8,883	3.153	1.070	-0.219	2.503	0.079	0.174	0.359	0.291	0.097
Cl.A4	80歳以上男性	100	4,700	3.284	0.992	-0.182	2.667	0.042	0.157	0.385	0.308	0.109
	無職	533	25,051	3.089	1.011	-0.201	2.610	0.072	0.188	0.388	0.285	0.068
	1人暮らし	298	14,006	3.096	0.998	-0.164	2.692	0.067	0.183	0.409	0.269	0.072
Cl.A5	70代	454	21,338	3.135	0.979	-0.242	2.724	0.061	0.174	0.400	0.300	0.065
	60代女性	224	10,528	3.169	0.974	-0.270	2.736	0.056	0.169	0.393	0.315	0.067
	70代男性	238	11,186	3.096	0.968	-0.244	2.691	0.061	0.187	0.400	0.298	0.054
	70代女性	213	10,011	3.182	0.985	-0.241	2.762	0.057	0.161	0.403	0.301	0.078
	専業主婦・主夫	303	14,241	3.229	1.018	-0.266	2.679	0.060	0.156	0.377	0.313	0.095
Cl.A6	全体	2,305	108,335	3.131	1.040	-0.228	2.548	0.074	0.181	0.366	0.298	0.081
	駿河区	706	33,182	3.184	1.016	-0.246	2.639	0.062	0.169	0.376	0.308	0.085
	女性	1,245	58,515	3.161	1.044	-0.250	2.560	0.073	0.173	0.362	0.306	0.087
	50代	422	19,834	3.111	1.022	-0.239	2.581	0.073	0.183	0.374	0.300	0.070
	60代	441	20,727	3.117	0.993	-0.245	2.654	0.065	0.182	0.388	0.301	0.064
	80歳以上	270	12,690	3.231	1.036	-0.244	2.623	0.062	0.158	0.371	0.305	0.104
	50代男性	191	8,977	3.094	0.995	-0.222	2.602	0.066	0.194	0.382	0.298	0.061
	50代女性	226	10,622	3.129	1.045	-0.257	2.564	0.079	0.172	0.368	0.301	0.080
	60代男性	213	10,011	3.056	1.004	-0.221	2.584	0.075	0.195	0.385	0.288	0.057
	80歳以上女性	169	7,943	3.210	1.053	-0.262	2.589	0.070	0.159	0.364	0.305	0.102
	バート・アルバート	346	16,262	3.136	1.009	-0.239	2.624	0.066	0.179	0.379	0.303	0.072
	家族と同居 (配偶者、子ども)	811	38,117	3.141	1.044	-0.256	2.527	0.075	0.180	0.355	0.311	0.080
	家族と同居 (配偶者のみ)	645	30,315	3.145	1.009	-0.252	2.615	0.065	0.179	0.373	0.310	0.072

に高めで低評価側の厚みが残る一方、尖度が高めで中心 (特に 3 付近) への集中が示唆される。属性としては「定年後＋独居」に近い層が含まれており、実務上は孤立・生活支援などの観点と関連づけて読み解く余地がある。

■ Cl.A5：高齢層 (中心集中・相対的にやや高評価寄り)

70 代, 60 代女性・70 代男女, 専業主婦・主夫が含まれる。尖度が高めで中心集中傾向を共有しつつ、Cl.A4 より歪度が低側で、高評価側への集中が相対的に強い。

■ Cl.A6：全体近傍 (中庸・全体に近い)

全体属性, 駿河区, 女性, 50～60 代など主要属性を含む。歪度・尖度が中程度で極端ではないが、「全体属性に近い」まとまりとして分離したと解釈できる。

以上より、 D_{JB-SD} による属性別クラスタリングでは、属性的に説明可能なまとまりが得られ、歪度・尖度を併用することでクラスタの意味づけが比較的容易であった。

5.3 他手法との比較

本節では、 D_{JB-SD} (図 3) で解釈した 6 クラスタを基準に、CMD および W_1 , 他 JB 距離指標でクラスタ構造がどのように再編されるかを手法ごとに整理する。ここでの差分は「良い/悪い」の序列ではなく、各手法が距離として強調する情報 (形状

差／水準差／累積比率差／標本規模差)の違いとして解釈する。

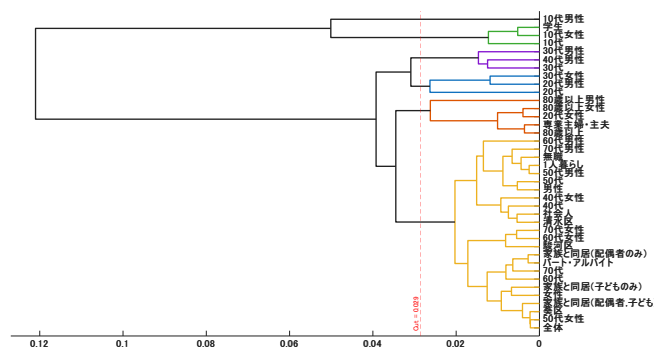


図 4 属性別デンドログラム (CMD, 群平均法, 6 クラス)

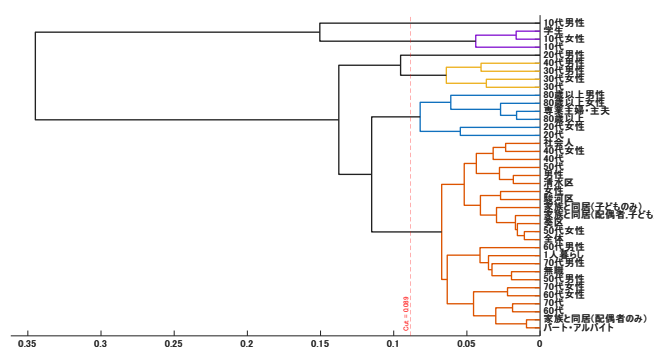


図 5 属性別デンドログラム (W_1 , 群平均法, 6 クラス)

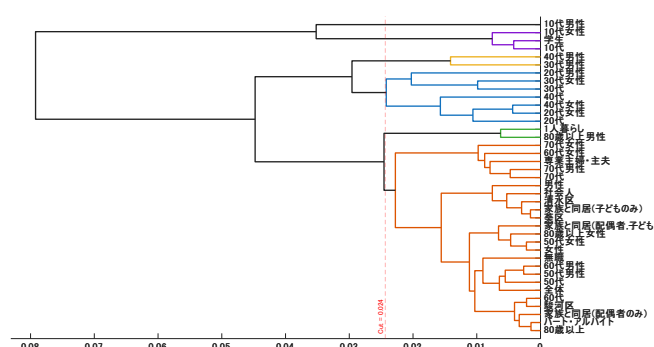


図 6 属性別デンドログラム (D_{JB-SI} , 群平均法, 6 クラス)

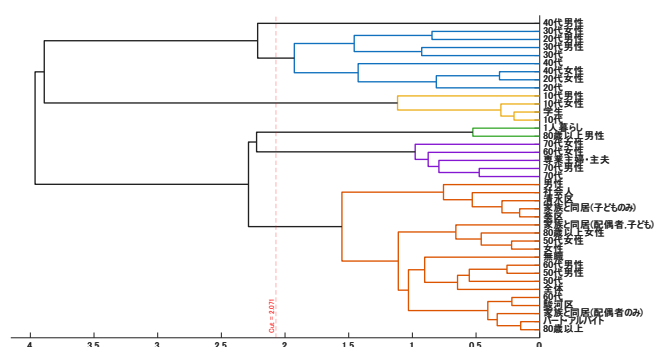


図7 属性別デンドログラム ($D_{JB-SD_{cap}}$, 群平均法, 6 クラス)

■ CMD：水準差・ばらつき差が効き、形状ベースの境界が再編される

図4はCMDによる属性別クラスタリング結果である。CMDは平均から高次中心モーメントまでの差を統合するため、歪度・尖度(形状差)に加えて平均水準やばらつき差も距離へ寄与し、 D_{JB-SD} で形成された「形状の近さ」による境界が再編され得る。例えば、現役世代クラスタ(CI.A1)から相対的に平均が低い30代・30代男性・40代男性が分離してまとまりを形成し、一方で平均が高い属性(20代女性, 80歳以上関連, 専業主婦・主夫)が年齢帯をまたいで同一クラスタへ統合される場面が見られる。また、若年・学生クラスタ(CI.A2)から最小属性(10代男性, $N=15$)が分離・孤立することがあり、小標本の偶然差がクラスタ構造に反映されやすい点に留意が必要である。したがってCMDは「水準差も含めた総合的な分布差」で整理したい場合に有用だが、「偏り・集中(歪度・尖度)を軸に説明したい」場合には D_{JB-SD} と異なるセグメント化を与え得る。

■ W_1 : 累積分布差 (段階比率差) が強調され, 端の比率差で孤立しやすい

図5は W_1 による結果である。1次元の W_1 は累積分布差を距離として反映するため、段階評価では段階比率(特に端の段階を含む比率差)がクラスタ形成要因として強く効く。その結果、10代男性(C1.A2)や20代男性(C1.A1)が単独で分離・孤立しやすい。前者は P_5 が相対的に大きく、後者は P_1 が相対的に大きいなど、特定段階の比率差が距離へ直接反映される帰結として理解できる。また、平均が高めの群(20代、20代女性、80歳以上関連、専業主婦・主夫)が再編される場面もあり、これは端の比率差だけでなく累積分布の全体差(分布中心のずれを含む)として近接した結果と解釈できる。したがって W_1 は分布差を素直に反映する一方、差異の要因(端の比率差/中心のずれ/小標本ゆえの偶然変動)を基本統計量と併読して判断する必要がある。

■ D_{JB-SI} : 標本サイズ補正がないため、小標本の形状差が前面に出て孤立しやすい

図 6 は D_{JB-SI} による結果である D_{JB-SI} は標本サイズ補正を持たないため、歪度・尖度の推定誤差 (特に小標本で大きい) がそのまま距離に反映され、形状が相対的に特異に見える属性が分離・孤立しやすい。例えば、若年・学生クラスタ (Cl.A2) から 10 代男性 ($N = 15$) が分離したり、現役世代内でも 30~40 代男性が枝分かれするなど、「形状差が出やすいが推定は不安定」な属性が距離構造を左右する場面が生じ得る。したがって D_{JB-SI} は形状差を強調して特異群候補を抽出する参照軸として有用だが、小標本属性の解釈には慎重さを要する。

■ $D_{JB-SDcap}$: 大標本同士 (または大標本と他群) の距離の過度な増大を抑え, 主要クラスタの統合を促し得る

図7は $D_{JB-SDcap}$ による結果である $D_{JB-SDcap}$ は有効標本サイズに上限を導入することで、大標本領域で距離が標本規模に応じて過度に増大する挙動を緩和する設計である。その結果、 D_{JB-SD} で独立していた大標本属性中心のクラスタ (代表的属性 Cl.A3) と、全体近傍クラスタ (Cl.A6) が統合されるなど、距離構造が相対的に「圧縮」され得る。一方で、主要ク

ラストが統合されると (クラスタ数を固定している条件では) 境界属性が押し出されて分離・孤立が生じる場合があるため、 D_{JB-SD} の結果に対して「大標本が構造を支配していないか」を点検する補助軸としての位置づけが考えられる。

5.4 考 察

本実験では、ウェルビーイングデータの属性別集計に対し、提案手法 (D_{JB-SD} , D_{JB-SI} , $D_{JB-SDcap}$) と比較手法 (CMD, W_1) で距離行列を構成し、階層クラスタリング (群平均法, 6 クラスタ固定) により実務上の解釈可能性と手法差を確認した。

D_{JB-SD} は歪度 (偏り) と尖度 (集中) という形状指標を軸に、属性的に説明しやすいクラスタ (例: 現役世代, 若年・学生, 高齢層, 全体近傍など) を与え, 小標本属性を含む状況でも標本サイズ補正により過敏な分離・孤立が生じにくい点が実務上の利点として整理できる。

一方で CMD は平均・ばらつき差を含む中心モーメント差を, W_1 は累積分布差 (段階比率差) を強調するため, D_{JB-SD} の形状ベースの境界が再編され得る。この差分は優劣ではなく, 「何を距離として強調するか」の違いとして理解でき, 実務では D_{JB-SD} を基準に, CMD・ W_1 を別視点の参照軸として重ねる補完運用が有効である。

JB 距離指標内部では, D_{JB-SI} は補正なしのため小標本の形状差が距離へ反映されやすく, $D_{JB-SDcap}$ は大標本領域での過度な距離増大を緩和する補助軸となる。したがって, D_{JB-SD} を基準に, D_{JB-SI} で特異群候補を点検し, $D_{JB-SDcap}$ で大標本支配の有無を点検する, という使い分けが自然である。

なお, 本節はクラスタ数を 6 に固定しているため, クラスタ数の最適化や孤立クラスタの扱いは今後の課題である。

6 おわりに

本研究では, 段階評価 (リッカート尺度) データにおける属性集団間の評価分布差を, 平均値などの単一指標では捉えにくい「分布形状差」として扱うため, 歪度・尖度に基づく非類似度指標 JB 距離を提案した。具体的には, 形状差の大きさを歪度差・尖度差の統合として定量化する D_{JB-SI} に加え, 標本規模差の影響を距離に織り込む D_{JB-SD} , および大標本域での過度な距離増大を抑える $D_{JB-SDcap}$ を定式化した。

シミュレーション (実験 1) では, D_{JB-SD} が小標本・不均衡条件で偽陽性率の上振れを相対的に抑えつつ, 多くの条件で比較手法と同等以上の検出力を示し, 形状差検出に有効であることを確認した。MovieLens (実験 2) では, 均衡条件で比較手法 (CMD, W_1) と高い整合性を示す一方, 不均衡条件では D_{JB-SD} の設計差が距離構造・クラスタ構造の差として顕在化し, 復元度 (擬似真ラベル ARI) の観点で $D_{JB-SD} \geq D_{JB-SI}$ となる場面も確認された。ウェルビーイング調査 (実験 3) では, D_{JB-SD} が歪度 (偏り)・尖度 (集中) に基づき解釈しやすい属性クラスタを与え, 小標本属性の過敏な孤立を相対的に抑え得ることを示した。以上の結果は, JB 距離が平均値などの単一の基本統計量だけでは捉えにくい評価傾向の違いを, 歪度 (偏

り)・尖度 (集中) という形状差として要約し, 距離行列とクラスタとして提示できることを意味する。

以上より, JB 距離は (i) 歪度・尖度という少数の形状指標に基づき解釈可能な分布差を与える点, (ii) 標本規模差を考慮する設計により, 小標本・不均衡を含む実務データでも距離構造を安定化し得る点を特徴とする非類似度指標として位置づけられる。

実務上は, (i) 属性別に「誰が似た傾向か」をセグメントとして整理し施策対象の説明に接続する, (ii) 項目別に「何が似た課題・強みか」を棚卸して論点整理や優先順位付けの起点を与える, (iii) 項目 × 属性別に「誰の・何の課題か」を特定して施策ターゲットの具体化へつなげる, といった単位での活用が想定される。ただし項目 × 属性の分割は小標本化を招き得るため, 標本規模と推定誤差に留意し, 必要に応じて CMD や W_1 を併用して「水準差を含む分布差」「段階比率差」の観点を補完することが望ましい。

今後の課題として, n_{cap} の適切な設定やクラスタ数の選定, ならびに多峰性など歪度・尖度だけでは捉えにくい分布差への対応が挙げられる。

文 献

- [1] F. Maxwell Harper and Joseph A. Konstan. The movielens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems*, Vol. 5, No. 4, pp. 19:1–19:19, 2015.
- [2] Lawrence Hubert and Phipps Arabie. Comparing partitions. *Journal of Classification*, Vol. 2, pp. 193–218, 1985.
- [3] Carlos M. Jarque and Anil K. Bera. A test for normality of observations and regression residuals. *International Statistical Review*, Vol. 55, No. 2, pp. 163–172, 1987.
- [4] Leonid V. Kantorovich. On the translocation of masses. *Doklady Akademii Nauk SSSR*, Vol. 37, No. 7–8, pp. 227–229, 1942.
- [5] Rensis Likert. A technique for the measurement of attitudes. Technical Report 140, The Science Press, 1932.
- [6] Nathan Mantel. The detection of disease clustering and a generalized regression approach. *Cancer Research*, Vol. 27, No. 2, pp. 209–220, 1967.
- [7] Cédric Villani. *Topics in Optimal Transportation*, Vol. 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
- [8] Werner Zellinger, Bernhard Moser, Thomas Grubinger, and Susanne Saminger-Platz. Central moment discrepancy (cmd) for domain-invariant representation learning. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2017.

意思決定を特徴づける形容詞的修飾を用いた 大規模言語モデルの実証的分析

馬田 光琉[†] 山本 修平^{††}

[†] 筑波大学 情報学群 知識情報・図書館学類 〒305-8550 茨城県つくば市春日 1 丁目 2 番地

^{††} 筑波大学 図書館情報メディア系 〒305-8550 茨城県つくば市春日 1 丁目 2 番地

E-mail: [†]ts2211866@u.tsukuba.ac.jp, ^{††}syamamoto@slis.tsukuba.ac.jp

あらまし 大規模言語モデル (LLM) を意思決定支援に応用するには, LLM の意思決定特性を明らかにし, 制御する手法の確立が重要である. 本研究では, リスク選好および時間選好を表す形容詞をプロンプトに含めることで, LLM の意思決定特性の変化を分析した. GPT-4o と Llama3.2 を対象とした実験の結果, GPT-4o ではリスク愛好的な形容詞により損失回避係数が上昇し, 現在志向な形容詞により時間割引率が上昇するなど, 形容詞の意味と一致する変化が観察された. 一方, Llama3.2 では形容詞による選好制御が限定的であった. 本研究は, 形容詞を用いたプロンプト設計により LLM の意思決定特性を制御できる可能性を示した.

キーワード LLM, 行動経済学, リスク選好, プロスペクト理論, 時間選好, 双曲割引モデル

1 はじめに

1.1 研究背景

近年, 大規模言語モデル (Large Language Models; LLM) は急速に発展し, 文章生成や要約, 対話応答, 情報検索支援, コード生成など多様なタスクにおいて活用されている. さらに, 意思決定のドメインにおいても LLM が活用される事例が存在する. 例えば, Chiang らは, 人間の集団意思決定において, 「悪魔の代弁者」と呼ばれる, 議論に反対意見を述べるような LLM を導入し, 集団意思決定プロセスがどのように変化するかを検証している [1]. また, Kitadai らは, 大規模言語モデル駆動型マルチエージェントシミュレーション (LLMs-driven MAS) を用いて最後通牒ゲームを行い, 人間の実験結果を再現する研究を行っている [2], [3].

このような研究動向の中で, LLM を意思決定支援や行動シミュレーションに応用するためには, モデルがどのような意思決定特性を内部的に備えているのかを明らかにすることが重要である [4]. すなわち, LLM が不確実性や遅延報酬を伴う選択課題に対してどのような判断傾向を示すかを明示的に測定し, 人間の行動に関する理論的枠組みとの対応関係を検討する必要がある.

行動経済学の分野では, 人間の意思決定に影響を与える心理的要因をリスク選好, 時間選好, 社会選好という三つの観点から分析することが一般的である [5]. 本研究では, 個人の意思決定において特に影響を与えるリスク選好と時間選好の二つに着目する [3]. リスク選好とは, 人間が不確実性を伴う選択に対してどの程度リスクを取るか, あるいは回避するかを示す指標である [6]. リスクを取る傾向が強い場合はリスク愛好的であり, 逆にリスクを避ける傾向が強い場合はリスク回避的とされる. リスク回避的な意思決定は, 安全性や確実性を重視する心理的傾向に基づくものであり, 一方でリスク愛好的な選択は, 報酬

の大きさや挑戦意欲を重視する心理的傾向に基づくものである. 一方, 時間選好とは, 人間が将来の利益を現在の利益よりもどの程度重視するかを示す指標であり [7], 将来の利益を重視する場合には割引率が低く, 逆に目先の利益を優先する場合には割引率が高くなる.

近年, これらの理論枠組みを用いて, LLM が人間の意思決定パターンをどの程度再現できるかを検証する研究が行われている [4], [8], [9]. 例えば, Jia らは, 商用の複数モデル (ChatGPT-4.0-Turbo, Claude-3-Opus, Gemini-1.0-pro) を対象に, 選択リスト形式で質問を与え, その回答を元に「リスク選好」「確率加重」「損失回避」という行動経済学の三つの要素をどの程度示すのかを測定している [4]. 同様に, 吉川らは, LLM に対してペルソナを与え, リスク選好や時間選好がどの程度表出されるかを測定している [8], [9].

しかし, 既存研究の多くは, LLM が元から備える選好構造を「観察的に測定する」段階にとどまっており, 特定の意思決定特性をモデルに対して意図的に付与し, その影響を体系的に評価する手法は十分に検討されていない.

1.2 研究目的

以上の背景から, 本研究では LLM の内部的な意思決定特性を行動経済学の理論枠組みに基づいて測定し, さらにプロンプト設計を通じて特性を付与・調整する方法を検討する. これにより, LLM が人間の意思決定特性をどの程度再現できるかを明らかにし, 意思決定モデルとしての LLM の特性を実証的に分析することを目的とする. この目的を達成する方法として, 本研究では Imasaka ら [10] の手法に着目した. 彼らは, 特性を表す形容詞をプロンプトに含めることで LLM の性格的傾向を制御できる可能性を示している. このような手法を応用すれば, リスク選好や時間選好を LLM に付与・調整することが理論的に可能であると考えられる.

本研究では実際に, GPT-4o と Llama3.2 を対象として, リ

スク選好および時間選好を表す形容詞をプロンプトに含めることで、選好がどのように変化するかをパラメータ推定によって分析した。条件として、選好を表す形容詞を一つ与えた場合と二つ与えた場合、選好の異なる形容詞を同時に与えた場合の三つの条件で実験を行った。その結果、GPT-4o ではリスク愛好的な形容詞により損失回避係数 λ が上昇し、現在志向の形容詞により時間割引率 k が上昇するなど、形容詞の意味と一致する方向へのパラメータの変化が観察された。また、同じ選好の形容詞を二つ与えても効果が累積しないこと、選好の異なる形容詞を同時に与えると片方の選好が優勢になったり、互いの選好を打ち消し合ったりすることも明らかになった。一方、Llama3.2 では多くの条件で、提示する選択肢を入れ替えた場合の回答の一致度が低く、形容詞の意味と矛盾した変化も観察され、形容詞に応じて選好を変化させる能力が LLM によって異なることが明らかになった。これらの結果から、意思決定特性を表す形容詞をプロンプトに含めることで LLM のリスク選好および時間選好を一定程度制御できることが示された。

2 関連研究

2.1 LLM の意思決定特性と行動経済学的分析

Jia ら [4] は、大規模言語モデルの意思決定特性を、行動経済学の理論に基づいて体系的に評価する枠組みを提案した。この研究では、リスク選好、損失回避、確率加重の三つの側面から LLM の判断傾向を分析し、ChatGPT-4.0-Turbo, Claude-3-Opus, Gemini-1.0-Pro の三つの商用モデルを対象に比較実験を行っている。実験では、行動経済学実験で用いられる選択リスト形式の課題を LLM に提示して得られた結果を、構築した TCN モデルを使用して、プロスペクト理論に基づくパラメータ（リスク回避度、確率加重、損失回避係数）を推定した。その結果、三つの LLM はいずれもリスク回避的な傾向を示していたが、リスク回避度・損失回避度・確率加重の度合いは LLM によって異なることが確認された。また、人口統計学的属性を付与した際の実験では、各 LLM の意思決定傾向に変化が観察された。その他にも、LLM の意思決定特性を行動経済学的視点から分析した研究が報告されている。吉川ら [8] は、LLM にペルソナとなる情報を与えた上で、プロスペクト理論に基づき、利得および損失領域におけるリスク選好の非対称性を検証し、LLM が損失回避的な傾向を示すことを明らかにした。さらに、同じ研究グループは、ペルソナを設定した LLM を対象に双曲割引モデルを用いて時間選好を分析し、報酬額や遅延時間に応じて割引率が変動する傾向を確認している [9]。

2.2 LLM を用いた意思決定エージェントに関する研究

Li ら [11] は、LLM が「言葉の上で示す考え方」と「エージェントとして実際にとる行動」との間に一貫性がないことを指摘した。彼らの研究では、ある LLM が社会的属性に関する質問に対して中立的で公平な回答をする一方で、社会的属性を付与したエージェントとして意思決定を行わせると、その行動選択に偏りが現れることが確認された。Echterhoff ら [12]

は、LLM が人間由来の学習データに起因して認知バイアスを示す可能性を指摘し、それを体系的に評価・緩和する枠組み「BIASBUSTER」を提案した。BIASBUSTER では、アンカリングやフレーミングなどの典型的な認知バイアスを定量的に測定し、さらに Self-Help 法により、モデル自身がプロンプトを再構成する仕組みを導入したことで、認知バイアスを低減した LLM 出力を実現することが可能となった。同様に、角田ら [13] は、LLM が人間の認知バイアスを受け継ぐことを指摘した。LLM が受け継いだ認知バイアスを軽減するため、クラウドソーシングで提案された社会的バイアス軽減手法を LLM のプロンプトに適用するアプローチを導入した。このアプローチによって、LLM の出力に含まれる認知バイアスの一部を軽減し、合理的な判断を行うようになることが示された。

2.3 本研究の位置づけ

2.1 節で述べた研究は、LLM が人間の意思決定バイアスを部分的に再現しつつも、LLM ごとに異なる特性を持つことを示している。さらに、人口統計学的属性を与えることで、LLM の意思決定傾向が変化することも明らかとなった。一方で、これらの研究は LLM にペルソナや条件を与えることでその意思決定傾向を観察する段階にとどまっており、LLM の意思決定特性を意図的に制御・操作する手法については十分に検討されていない。また、2.2 節で述べた研究は、LLM に内在するバイアスを検出したり、その影響を最小化することを目的としている。しかし、本研究では、そのようなバイアスを単に欠陥として除去するのではなく、人間が持つ特性や判断の偏りをあえて LLM に付与し、そのようなエージェントによる意思決定を再現・分析することを目的とする。

3 分析手法

3.1 手法概要

本研究では、大規模言語モデル（LLM）に対して認知的特性を表す形容詞をプロンプト内に付与し、その影響が意思決定特性（リスク選好・時間選好）にどのような変化を与えるかを観察する手法を提案する。本研究で提案する分析手法の概要を図 1 に示す。手法は大きく Step1 と Step2 から構成される。

Step1 では、金額と確率（もしくは期間）のリストおよび認知的特性を表す形容詞を用いて選択課題を生成し、LLM に複数回答させることで選択課題タスクの結果を得る。また、選択肢の順序による回答への影響を考慮して、選択肢を入れ替えた場合も含めて合計 40 回、LLM に回答させる。

Step2 では、Step1 で得られた選択データに対してプロスペクト理論および双曲割引モデルを用いて最尤推定を行い、意思決定特性を表すパラメータを推定する。これにより、形容詞の付与によって LLM の意思決定傾向がどの程度変動するかを観察する。

今回主に使用したモデルは、gpt-4o-2024-08-06 および llama3.2:latest であり、両モデルにおいて temperature は

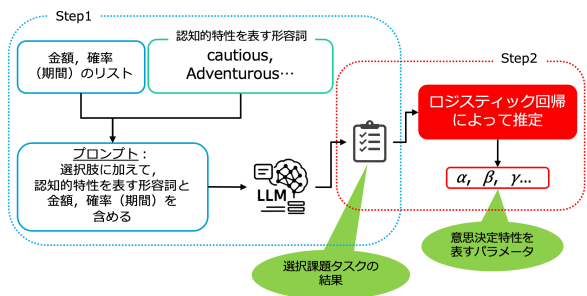


図 1 本研究で提案する分析手法の概要図

0 に設定している。¹これ以降、gpt-4o-2024-08-06 および llama3.2:latest はそれぞれ GPT-4o および Llama3.2 と略記する。

3.2 Step1：選択課題データ生成

Step1 では、意思決定特性としてリスク選好と時間選好を取り上げ、それらを表す形容詞を用いたプロンプトを設計する。続いて、リスク選好タスクおよび時間選好タスクの選択課題を自動生成し、プロンプトとして LLM に複数回回答させることで選択課題データを取得する。

3.2.1 本研究で扱う意思決定特性の概要

本研究で対象とする意思決定特性は、リスク選好 [6] と時間選好 [7] の二種類である。リスク選好は、人間がリスクをどの程度好むか、あるいは避けるかを表す特性である。本研究では、リスク選好の水準を「リスク回避的」「中立」「リスク愛好的」の三水準として扱う。時間選好は、人間が現在の利益を将来の利益よりもどの程度重視するかを表す特性である。本研究では時間選好の水準を「現在志向」「中立」「将来志向」の三水準として扱う。

3.2.2 プロンプト設計

ここでは、Imasaka らの手法 [10] を参考に、リスク選好および時間選好を LLM に付与する。Imasaka らは、大規模言語モデル (LLM) の性格特性が検索タスクにおけるクエリ生成行動へ与える影響を分析した。この研究では、ビッグファイブ理論に基づき、協調性 (Agreeableness)、誠実性 (Conscientiousness)、外向性 (Extraversion)、神経症傾向 (Neuroticism)、開放性 (Openness) の五つの性格特性を LLM に付与した上で、その出力傾向を評価している。実験では、ビッグファイブ理論に基づく各性格特性を表す形容詞を多数含むプロンプトを LLM に提示し、それぞれの性格特性を付与したエージェントを生成した。その結果、性格特性を付与した LLM は、プロンプトに与えられた形容詞の修飾に応じて生成されるクエリの長さや語彙の多様性が変化することが確認された。この Imasaka らの手法を参考に、認知的特性を表す形容詞を含めたプロンプトを、吉川らが用いたプロンプトテンプレート [8], [9] に組み込むことで、LLM にリスク選好および時間選好を付与するプロンプト

1：この他にも gpt-3.5-turbo-0125, llama3.1:latest, llama3.3:latest, phi4:latest, gemma3:27b, mistral:latest を使用して実験を行ったが、詳細な結果分析は gpt-4o-2024-08-06 および llama3.2:latest に絞って行う。

表 1 本研究で用いた認知的特性を表す形容詞

意思決定特性	選好	形容詞
リスク選好・時間選好	中立	Moderate (中庸的な)
		Balanced (均衡をとった)
	リスク愛好的	Adventurous (冒険的な)
		Audacious (大胆な)
リスク選好	リスク回避的	Cautious (慎重な)
		Conservative (保守的な)
時間選好	将来志向	Patient (忍耐強い)
		Prudent (慎重な)
	現在志向	Impatient (せっかちな)
		Impulsive (衝動的な)

設計を行う。本研究で用いる形容詞と、リスク選好タスクと時間選好タスクで提示するプロンプトをそれぞれ以下に示す。

リスク選好タスクのプロンプトテンプレート

You have a behavioral economic characteristic of being {neutral or low or high} in risk aversion. To explain in detail, you are {認知的特性を表す形容詞}.

I will now ask you multiple-choice questions, so please output only the letter of your answer A or B, and nothing else.

Question: There are two ways to {receive or lose} the prize money. Based on your intuition and personal risk preferences, which of the following options would you choose?

A. {receive or lose} {金額 A} for certain.

B. {確率 p_1 } chance to {receive or lose} {金額 B_1 } or {確率 p_2 } chance to {receive or lose} {金額 B_2 }.

Answer(A or B):

時間選好タスクのプロンプトテンプレート

You have a behavioral economic characteristic of having a {neutral or low or high} time discount rate. To explain in detail, you are {認知的特性を表す形容詞}.

I will ask you multiple-choice questions, please answer with either SSR or LLR, and nothing else.

Question: You have won the lottery. There are two ways to receive the prize money. Based on your intuition and personal time preferences, which of the following options would you choose?

A. Receive {金額 A} after 0 days.

B. Wait until {期間 t } days later and receive {金額 $B(> A)$ }.

Answer:

これらのテンプレートを用いて、使用する LLM に意思決定特性を与えた上で、リスク選好タスクでは確実に得られる、もしくは失う金額 A と不確実に得られる、もしくは失う金額 B_1, B_2 およびその確率 p_1, p_2 の組み合わせを変化させた選択課題を生成する。同様に、時間選好タスクでは即時に得られる金額 A と遅延して得られる金額 B およびその遅延期間 t の組

み合わせを変化させた選択課題を生成する。また、意思決定特性の与え方として、形容詞の一つ用いる方法、同じ選好を表す形容詞を二つ用いる方法、選好の異なる形容詞を同時に用いる方法の三通りを検討する。

3.2.3 選択課題タスク

選択課題タスクの生成では 3.2.2 節で示したプロンプトテンプレートに基づき、Python によりリスク選好タスクおよび時間選好タスクの選択課題データを自動生成している。

まず、リスク選好タスクについて説明する。本研究におけるリスク選好タスクは、確実に得られる、もしくは失う金額 A と不確実に得られる、もしくは失う二つの金額 B_1, B_2 およびその確率 p_1, p_2 の組み合わせを変化させた選択課題を生成し、LLM がどちらの選択を行うかを観察するものである。ここで用いる金額リストと確率リスト [8] は以下の通りである。

- 金額リスト (単位: \$):
[100, 500, 750, 1000, 2500, 5000, 7500, 10000]
- 確率リスト:
[0.01, 0.05, 0.10, 0.20, 0.25, 0.30, 0.40, 0.50, 0.60, 0.75, 0.80, 0.90, 0.95, 0.99]

これらのリストから、不確実な選択肢 B に対応する二つの金額 B_1, B_2 と確率 p_1 , および p_2 ($= 1 - p_1$) を抽出する。次に、確実な選択肢 A の金額 A は、選択肢 B の期待値

$$EV_B = p_1 \times B_1 + p_2 \times B_2$$

に 1.0~1.01 の範囲の乱数係数を掛けて決定している。このようにして、金額が得られる選択課題タスク、金額を失う選択課題タスクをそれぞれ 10 問ずつ、合計 20 問生成し、選択肢の順序を入れ替えたものも含めて合計 40 問のリスク選好タスクがプロンプトテンプレート (3.2.2 節) に挿入され、LLM に提示される。

同様に、時間選好タスクについても、Python により選択課題データを自動生成している。本研究における時間選好タスクは、即時に得られる金額 (SSR: Short-Soon Reward) と遅延して得られる金額 (LLR: Long-Late Reward) およびその遅延期間の組み合わせを変化させた選択課題を生成し、LLM がどちらの選択を行うかを観察するものである。ここで用いる金額リストと遅延期間リスト [9] は以下の通りである。

- 即時報酬リスト (単位: \$):
[100, 500, 1000, 5000, 10000]
- 報酬比率リスト:
[1.2, 1.5, 2.0]
- 遅延期間リスト (単位: day):
[1, 7, 30, 180, 365]

これらのリストから、即時報酬 A 、報酬比率 r 、遅延期間 t を抽出し、遅延報酬 B を

$$B = r \times A$$

として計算することで、選択課題タスクを生成する。このようにして、選択課題タスクを 20 問生成し、選択肢の順序を入れ替えたものも含めて合計 40 問の時間選好タスクがプロンプトテンプレート (3.2.2 節) に挿入され、LLM に提示される。

3.3 Step2: パラメータ推定

Step1 で得られたリスク選好タスクおよび時間選好タスクの選択データに対して、Step2 ではプロスペクト理論の価値関数、確率加重関数および双曲割引モデルを用いて、LLM の意思決定特性を表すパラメータを最尤推定する。推定されたパラメータを比較することで、形容詞の付与によって LLM の意思決定傾向がどの程度変動するかを観察する。

3.3.1 プロスペクト理論によるリスク選好の推定

リスク選好タスクに対しては、プロスペクト理論 [14] に基づき、利得と損失に対する価値関数と確率加重関数を用いたモデルを採用する。金額 x に対する価値関数 $v(x)$ は、利得と損失で異なるべき乗型の関数として次のように定義する。

$$v(x) = \begin{cases} x^\alpha & (x \geq 0) \\ -\lambda(-x)^\beta & (x < 0) \end{cases} \quad (1)$$

ここで、 α は利得の感応度、 β は損失の感応度、 λ は損失回避係数を表すパラメータである。一般に $0 < \alpha, \beta \leq 1$ のとき、利得・損失のいずれの領域においても、金額の絶対値が大きくなるほど主観的価値の増加幅が小さくなるという限界効用逓減 [15] の性質を示す。また、 $\lambda > 1$ のとき、損失は同額の利得よりも大きく評価される。

確率 $p \in [0, 1]$ に対しては、確率加重関数を用いて、単一のパラメータ γ をもつ次の関数を用いる。

$$w(p) = \frac{p^\gamma}{\{p^\gamma + (1-p)^\gamma\}^{1/\gamma}} \quad (2)$$

$\gamma < 1$ のとき、実際に発生する確率が小さい事象の確率が過大評価され、実際に発生する確率が大きい事象の確率が過小評価される。一方、 $\gamma = 1$ のとき $w(p) = p$ となり、実際に発生する確率と一致する。

3.2.2 節のプロンプトテンプレートで示した通り、本研究のリスク選好タスクでは、確実に金額 A を得る選択肢 A と、確率 p_1 で金額 B_1 、確率 p_2 で金額 B_2 を得る選択肢 B の二択が提示される。40 回の回答のうち、 i 回目に提示される金額と確率 (A, B_1, B_2, p_1, p_2) をそれぞれ、

$$q_i = (A_i, B_i^+, B_i^-, p_i^+, p_i^-) \quad (3)$$

と書く。このとき、人間の金額に対する主観的価値を表す効用関数 $U_A(q_i), U_B(q_i)$ は

$$U_A(q_i) = v(A_i) \quad (4)$$

$$U_B(q_i) = w(p_i^+) v(B_i^+) + w(p_i^-) v(B_i^-) \quad (5)$$

と表される。

次に、LLM の選択データと効用関数との関係を定式化する。このとき、 i 回目の質問において、選択肢 A が選ばれる確率を μ_i とおき、次のロジスティックモデルで定義する。

$$\mu_i = \frac{1}{1 + \exp\{-\tau(U_A(q_i) - U_B(q_i))\}} \quad (6)$$

μ_i は選択肢 A と B の主観的価値の差に応じて変化すると考える。一般に、 $U_A(q_i)$ が $U_B(q_i)$ より大きいほど選択肢 A が選ばれやすく、両者がほぼ同程度であれば選択がランダムになると想定できる。ここで $\tau > 0$ はロジスティックモデルの傾きを表すパラメータであり、 τ が大きいほど主観的価値の差が過剰に評価され、 μ_i は 0 か 1 のどちらかにより近づく。一方、 τ が小さいほど主観的価値の差はそれほど評価されなくなり、選択がランダム ($\mu_i \approx 0.5$) に近づく。

質問回数 $i = 1, \dots, N$ において LLM が選択肢 A を選んだ場合を $y_i = 1$ 、選択肢 B を選んだ場合を $y_i = 0$ と定義し、このように定義した μ_i をパラメータとするベルヌーイ分布に従うと仮定すると、

$$y_i \sim \text{Bern}(\mu_i) \quad (7)$$

が成り立つ。したがって、 N 回の質問に対する回答系列 $y_1, \dots, y_N$ が得られる確率は

$$P(y_1, \dots, y_N) = \prod_{i=1}^N \mu_i^{y_i} (1 - \mu_i)^{1-y_i} \quad (8)$$

となる。

パラメータ集合 $\theta = (\alpha, \beta, \lambda, \gamma, \tau)$ に対する対数尤度は

$$\log L(y_1, \dots, y_N | \theta) = \sum_{i=1}^N \{y_i \log \mu_i + (1 - y_i) \log(1 - \mu_i)\} \quad (9)$$

と書ける。本研究では、この対数尤度を最大化する制約付き最適化問題を解くことでパラメータ θ を推定する。

$$\hat{\theta} = \arg \max_{\theta} \log L(y_1, \dots, y_N | \theta) \quad (10)$$

$$\text{s.t. } 0 < \alpha, 0 < \beta, 1 \leq \lambda, 0 \leq \gamma \leq 1, 0 < \tau$$

ここで $\hat{\theta}$ を最尤推定値として用いる。最尤推定では、Python の数値最適化ライブラリである `scipy.optimize` を用いて、負の対数尤度 $-\log L$ を最小化する形で $\hat{\theta} = (\alpha, \beta, \lambda, \gamma, \tau)$ を求める。推定された $\hat{\theta}$ のうち、 $\alpha, \beta, \lambda, \gamma$ を比較することで、プロンプトに付与した形容詞が LLM の意思決定特性に与える影響を分析する。

3.3.2 双曲割引モデルによる時間選好の推定

時間選好タスクに対しては、遅延報酬の主観的価値が時間とともに減衰すると仮定する双曲割引モデル [7] を用いる。金額 x を遅延時間 t 日後に受け取る時の主観的価値 $V(x, t)$ を

$$V(x, t) = \frac{x}{1 + kt} \quad (11)$$

と定義する。ここで、 $k \geq 0$ は割引率を表すパラメータであり、 k が大きいほど将来の報酬を強く割り引き、現在志向的な時間

選好を示す。

3.2.2 節のプロンプトテンプレートで示した通り、本研究の時間選好タスクでは、即時に金額 A を受け取る選択肢 SSR と、 t 日後に金額 B を受け取る選択肢 LLR の二択が提示される。40 回の回答のうち、 i 回目に提示される即時報酬・遅延報酬および遅延期間 (A, B, t) を、それぞれ $x_i^{\text{SSR}}, x_i^{\text{LLR}}, t_i^{\text{LLR}}$ と書き、即時報酬の受け取り時点は $t_i^{\text{SSR}} = 0$ とおく。このときのパラメータ集合を

$$q_i = (x_i^{\text{SSR}}, x_i^{\text{LLR}}, t_i^{\text{SSR}}, t_i^{\text{LLR}}) \quad (12)$$

と表す。

双曲割引モデルに基づく各選択肢の主観的価値は、式 (11) を用いて

$$V_{\text{SSR}}(q_i) = V(x_i^{\text{SSR}}, t_i^{\text{SSR}}) = x_i^{\text{SSR}} \quad (13)$$

$$V_{\text{LLR}}(q_i) = V(x_i^{\text{LLR}}, t_i^{\text{LLR}}) = \frac{x_i^{\text{LLR}}}{1 + kt_i^{\text{LLR}}} \quad (14)$$

と表される。

以降は、リスク選好タスクと同様の手続きでパラメータを推定した。すなわち、 i 回目の質問において即時報酬 (SSR) を選択する確率 μ_i をロジスティックモデルで定義し、ベルヌーイ分布に基づく対数尤度を最大化する制約付き最適化問題を解くことで、パラメータ $\theta = (k, \tau)$ を推定する。推定されたパラメータのうち、割引率 k を比較することで、プロンプトに付与した時間選好に関する形容詞が LLM の意思決定特性に与える影響の違いを評価する。

4 リスク選好に関する実験結果と考察

本章では、GPT-4o および Llama3.2 に対して、形容詞を用いた場合のリスク選好のパラメータ推定の結果を示すとともに、その考察を行う。なお、一致度は選択肢の順序を入れ替えても同じ選択肢を選んだ割合を示し、有効回答率はモデルが有効な回答を返した割合を示す。ここでいう有効な回答とは、「A」「B」のいずれかを選択する回答を指す。

4.1 動いたパラメータから考えられること

表 2 に、形容詞の一つ用いた場合のパラメータ推定の結果を示す。表 2 において、GPT-4o にリスク愛好的な形容詞 (Adventurous/Audacious) を与えた際に λ が 1.000 から 5.000 へ上昇した。 $\lambda = 5.000$ への上昇は、モデルが損失回避的な傾向を強めたことを示している。これは一見、リスク愛好的な形容詞の意味と矛盾するように見える。しかし、リスク愛好的な行動は「利得領域でリスクを取る」とこと「損失領域で損失を強く回避する」ことの両方を含みうる。本実験の結果は、GPT-4o がリスク愛好的な形容詞を後者の意味で解釈した可能性を示唆しており、具体的には、損失領域の選択課題タスクにおいて、たとえリスクを取ってでも損失額を減らしたいという思考が働いたと考えられる。また表 2 において、GPT-4o にリスク愛好的な形容詞を与えると、 α が 0.993 から 1.058 へ増加し、利得額の増加に対して主観的価値がより大きく増加するよ

表 2 形容詞を一つ用いた場合のパラメータ推定の結果

リスク選好	形容詞	gpt-4o-2024-08-06						llama3.2:latest					
		α	β	λ	γ	負の対数尤度	一致度	α	β	λ	γ	負の対数尤度	一致度
-	なし	0.993	1.001	1.000	0.997	43.371	85%	1.002	1.004	1.000	1.000	42.032	0%
中立	Moderate	0.968	1.017	1.000	1.000	38.434	90%	1.002	1.004	1.000	1.000	42.032	0%
	Balanced	0.968	1.111	1.000	1.000	42.815	95%	1.002	1.004	1.000	1.000	42.032	0%
リスク愛好的	Adventurous	1.058	0.977	5.000	1.000	19.956	100%	1.002	0.998	5.000	1.000	20.834	55%
	Audacious	1.058	0.977	5.000	1.000	19.956	100%	1.002	1.004	1.000	1.000	42.032	0%
リスク回避的	Cautious	0.968	1.110	1.000	1.000	44.210	100%	1.002	1.004	1.000	1.000	42.032	0%
	Conservative	0.968	1.110	1.000	1.000	44.210	100%	1.002	1.004	1.000	1.000	42.032	0%
-	人間の例 [16]	0.859	0.826	1.576	0.605	-	-	0.859	0.826	1.576	0.605	-	-

注) いずれの実験結果も有効回答率は 100%であった。

表 3 形容詞を二つ用いた場合のパラメータ推定の結果

リスク選好	形容詞	gpt-4o-2024-08-06						llama3.2:latest					
		α	β	λ	γ	負の対数尤度	一致度	α	β	λ	γ	負の対数尤度	一致度
-	なし	0.993	1.001	1.000	0.997	43.371	85%	1.002	1.004	1.000	1.000	42.032	0%
中立	Moderate + Balanced	0.968	1.111	1.000	1.000	42.815	95%	1.002	1.004	1.000	1.000	42.032	0%
リスク愛好的	Adventurous + Audacious	1.058	0.977	5.000	1.000	19.956	100%	1.002	0.999	1.000	1.000	26.981	35%
リスク回避的	Cautious + Conservative	0.968	1.110	1.000	1.000	44.210	100%	1.002	1.004	1.000	1.000	42.032	0%

注) いずれの実験結果も有効回答率は 100%であった。

表 4 選好の異なる形容詞を同時に用いた場合のパラメータ推定の結果

形容詞	gpt-4o-2024-08-06						llama3.2:latest					
	α	β	λ	γ	負の対数尤度	一致度	α	β	λ	γ	負の対数尤度	一致度
なし	0.993	1.001	1.000	0.997	43.371	85%	1.002	1.004	1.000	1.000	42.032	0%
Adventurous + Cautious	0.968	1.002	1.000	1.000	25.967	85%	1.002	1.004	1.000	1.000	27.904	15%
Adventurous + Conservative	0.950	0.995	1.000	0.988	17.324	75%	1.002	1.004	1.000	1.000	41.365	5%
Audacious + Cautious	0.968	1.003	1.000	1.000	26.288	70%	1.002	1.004	1.000	1.000	41.365	5%
Audacious + Conservative	0.968	1.004	1.000	1.000	25.198	55%	1.002	1.004	1.000	1.000	42.032	0%

注) いずれの実験結果も有効回答率は 100%であった。

うになったことを表している。また、 β が 1.001 から 0.977 へ低下し、損失額が増加しても主観的な損失感があまり増えないようになったことを表している。一方、表 2 において、中立な形容詞 (Moderate/Balanced) およびリスク回避的な形容詞 (Cautious/Conservative) を与えた場合は $\lambda = 1.000$ が維持された。これは $\lambda \geq 1$ という制約があるため、リスク回避的な形容詞を与えることによってさらに低下する余地がなかった。ただし、 β は 1.017 ~ 1.111 へ上昇しており、損失額の増加に対して主観的な損失感がより大きく増加するようになったことを意味する。 γ はほぼ全ての条件で 1.000 を維持し、形容詞を与えても確率の主観的評価には影響を与えにくいことが示された。

表 2 において、Llama3.2 では、Adventurous 単独の場合を除き、ほぼ全ての条件でパラメータが形容詞を与えない場合 ($\alpha \approx 1$, $\beta \approx 1$, $\lambda = 1$, $\gamma = 1$) と変化がなかった。これらの値はリスク中立的な選好を表しており、Llama3.2 が形容詞に応じて選好を変化させる能力が限定的であることを示している。唯一 Adventurous で $\lambda = 5.000$ が推定されたことは、この形容詞が Llama3.2 にとって特に強い影響を及ぼした可能性がある。

4.2 形容詞を一つ用いた場合と二つ用いた場合の違い

表 3 に形容詞を二つ用いた場合のパラメータ推定の結果を示す。表 2 と表 3 を比較すると、GPT-4o において、形容詞を一つ用いた場合と二つ用いた場合 (同一選好の形容詞を二つ組み合わせた場合) では、推定されるパラメータに大きな差は見られなかった。例えば、表 2 の Adventurous 単独の場合と表 3 の Adventurous + Audacious の場合で、いずれも $\alpha = 1.058$, $\beta = 0.977$, $\lambda = 5.000$, $\gamma = 1.000$ が推定された。このことから、同一選好の形容詞を重ねても効果が累積しないことが考えられる。

表 4 に、選好の異なる形容詞を同時に用いた場合のパラメータ推定の結果を示す。表 4 において、選好の異なる形容詞を組み合わせた場合では、リスク愛好的な形容詞とリスク回避的な形容詞を組み合わせても $\lambda = 1.000$ が維持され、 α は 0.950 ~ 0.968, β は 0.995 ~ 1.004 の範囲に収まった。このことから、相反する選好の形容詞が相互に打ち消し合い、結果として中立に近い選好が表現されたと解釈できる。

Llama3.2 では、表 2, 3, 4 のいずれにおいても、Adventurous 単独の場合を除きパラメータが形容詞を与えない場合の値

付近に留まった。形容詞の数や組合せによる系統的な差は観察されなかった。

4.3 人間との比較

人間の推定値の一例 ($\alpha = 0.859$, $\beta = 0.826$, $\lambda = 1.576$, $\gamma = 0.605$) [16] を参考として、本実験の結果と比較する。この被験者は、 $\alpha = 0.859$ および $\beta = 0.826$ いずれも 1 未満であり、利得・損失の両領域で限界効用逓減の特徴を持つ。また、 $\lambda = 1.576$ で損失回避的（損失を利得の約 1.6 倍重く評価する）、 $\gamma = 0.605$ で確率加重バイアスを持つという特徴を示している。表 2 において、形容詞を与えない場合、両モデルは $\alpha \approx 1$, $\beta \approx 1$, $\lambda = 1.000$, $\gamma \approx 1$ であり、これは、この被験者が限界効用逓減や損失回避といった特徴を持つのに対し、両モデルはリスク中立的な選好を示している点で異なる。表 2 において、リスク愛好的な形容詞を与えた場合、GPT-4o では $\alpha = 1.058$, $\beta = 0.977$, $\lambda = 5.000$ となった。 $\lambda = 5.000$ はこの被験者 ($\lambda = 1.576$) よりも大きく、損失を利得の 5 倍重く評価するという強い損失回避を示している。また、この被験者は $\alpha = 0.859$, $\beta = 0.826$ と両領域で限界効用逓減の特徴を持っているのに対して、GPT-4o では $\alpha > 1$ で利得に対する感応度が高く、 $\beta < 1$ で損失に対する感応度が低いという、人間とは異なる特徴を示している。表 2 においてリスク回避的な形容詞を与えた場合、GPT-4o では $\beta = 1.110$ となった。 $\beta > 1$ は損失に対する感応度が高いことを示し、損失額の増加に対して主観的な損失感がより大きく増加することを意味する。これに対し、この被験者は $\beta = 0.826 < 1$ であり、損失に対する感応度が低いことを示している。 γ については、表 2 において、いずれの条件でも $\gamma \approx 1.000$ であり、確率に対する主観的価値は客観的価値と一致している。この被験者は $\gamma = 0.605$ であり、小さな確率を過大評価し大きな確率を過小評価する傾向を持つが、両モデルではこのような確率加重バイアスは観察されなかった。

5 時間選好に関する実験結果と考察

本章では、GPT-4o および Llama3.2 に対して、形容詞を用いた場合の時間選好のパラメータ推定の結果を示すとともに、その考察を行う。なお、一致度は選択肢の順序を入れ替えても同じ選択肢を選んだ割合を示し、有効回答率はモデルが有効な回答を返した割合を示す。ここでいう有効な回答とは、「SSR」「LLR」のいずれかを選択する回答を指す。

5.1 動いたパラメータから考えられること

表 5 に、形容詞を一つ用いた場合のパラメータ推定の結果を示す。表 5 において、GPT-4o に現在志向の形容詞 (Impatient/Impulsive) を与えた際に k が 0.010 から 3.666 へ上昇した。 $k = 3.666$ への上昇は、モデルが「今すぐ報酬を得たい」という傾向を強めたことを示しており、Impatient (せっかちな) や Impulsive (衝動的な) という形容詞の意味と一致する変化である。一方、表 5 において、中立な形容詞 (Moderate/Balanced) および将来志向の形容詞 (Prudent/Patient) を与えた場合は $k = 0.010$ が維持された。これは形容詞を与え

ない場合にすでに $k = 0.010$ であったため、将来志向の形容詞によるさらなる低下の余地がなかったと考えられる。これは、GPT-4o のデフォルトの時間選好がすでに強い将来志向であることを示唆している。

表 5 において、Llama3.2 では、形容詞を与えない場合で $k = 3.664$ であり、GPT-4o とは異なり、デフォルトの時間選好が現在志向であった。また、中立の形容詞 (Moderate/Balanced) を与えた場合でも $k = 3.664$ が維持された。将来志向の形容詞 (Prudent/Patient) を与えた場合は k が 0.069 ~ 0.500 へ低下し、将来志向への変化が見られた。これは、形容詞の意味と一致する方向への変化であり、Llama3.2 が将来志向の形容詞を適切に解釈できたことを示している。しかし、Llama3.2 では現在志向の形容詞 (Impatient/Impulsive) を与えた場合に $k = 0.500$ となり、形容詞を与えない場合 ($k = 3.664$) よりむしろ低下するという逆転現象が観察された。これは形容詞の意味と矛盾する変化であり、Llama3.2 が現在志向の形容詞を正しく解釈できていなかった可能性がある。

5.2 形容詞を一つ用いた場合と二つ用いた場合の違い

表 6 に、形容詞を二つ用いた場合のパラメータ推定の結果を示す。表 5 と表 6 を比較すると、GPT-4o では、形容詞を一つ用いた場合と二つ用いた場合（同一選好の形容詞を二つ組み合わせた場合）で推定される k に差は見られなかった。中立・将来志向の形容詞では $k = 0.010$ 、現在志向の形容詞では $k = 3.666$ となり、形容詞の数によってより強い選好は現れなかった。

表 7 に、選好の異なる形容詞を同時に用いた場合のパラメータ推定の結果を示す。表 7 において、選好の異なる形容詞を組み合わせた場合では、Patient + Impatient および Patient + Impulsive で $k = 0.010$ 、Prudent + Impatient で $k = 3.666$ 、Prudent + Impulsive で $k = 0.500$ となった。Patient + Impatient および Patient + Impulsive では将来志向の形容詞 (Patient) が優勢に働き、Prudent + Impatient では現在志向の形容詞 (Impatient) が優勢に働いた。Prudent + Impulsive では $k = 0.500$ と中間的な値となり、両者の選好が打ち消し合い、結果として中立に近い選好が表現されたと解釈できる。

表 6 において、Llama3.2 では、形容詞を二つ用いたときに形容詞を与えない場合 ($k = 3.664$) から大きく変化する傾向が見られた。中立の形容詞を二つ与えた場合は $k = 0.501$ 、将来志向の形容詞を二つ与えた場合は $k = 0.010$ 、現在志向の形容詞を二つ与えた場合は $k = 0.500$ となった。いずれの場合も形容詞を一つ与えたとき以上に将来志向的な値へと変化しており、形容詞を二つ与えること自体が応答パターンに影響を与えている可能性がある。表 7 において、選好の異なる形容詞を組み合わせた場合では、Patient + Impatient, Patient + Impulsive, Prudent + Impatient で $k = 0.010$ 、Prudent + Impulsive で $k = 0.500$ となった。表 5 において、形容詞単体では、Patient で $k = 0.500$ 、Prudent で $k = 0.069$ 、Impatient および Impulsive で $k = 0.500$ であったが、表 7 の Patient + Impatient の組み合わせでは $k = 0.010$ となり、Patient 単体 ($k = 0.500$) や Impatient 単体 ($k = 0.500$) よりもさらに

表 5 形容詞を一つ用いた場合のパラメータ推定の結果

時間選好	形容詞	gpt-4o-2024-08-06			llama3.2:latest		
		k	負の対数尤度	一致度	k	負の対数尤度	一致度
-	なし	0.010	90.851	50%	3.664	20.723	95%
中立	Moderate	0.010	379.983	100%	3.664	20.723	95%
	Balanced	0.010	379.983	100%	3.664	20.723	95%
将来志向	Prudent	0.010	379.983	100%	0.069	87.842	55%
	Patient	0.010	379.983	100%	0.500	85.229	60%
現在志向	Impatient	3.666	0.000	100%	0.500	23.059	75%
	Impulsive	3.666	0.000	100%	0.500	23.059	75%
-	人間の例 [17]	2.6 ± 0.94	-	-	2.6 ± 0.94	-	-

注) いずれの実験結果も有効回答率は 100%であった。

表 6 形容詞を二つ用いた場合のパラメータ推定の結果

時間選好	形容詞	gpt-4o-2024-08-06			llama3.2:latest		
		k	負の対数尤度	一致度	k	負の対数尤度	一致度
-	なし	0.010	90.851	50%	3.664	20.723	95%
中立	Moderate + Balanced	0.010	379.983	100%	0.501	44.247	80%
将来志向	Prudent + Patient	0.010	379.983	100%	0.010	229.475	15%
現在志向	Impatient + Impulsive	3.666	0.000	100%	0.500	24.882	85%

注) いずれの実験結果も有効回答率は 100%であった。

表 7 選好の異なる形容詞を同時に用いた場合のパラメータ推定の結果

形容詞	gpt-4o-2024-08-06			llama3.2:latest		
	k	負の対数尤度	一致度	k	負の対数尤度	一致度
なし	0.010	90.851	50%	3.664	20.723	95%
Patient + Impatient	0.010	206.224	40%	0.010	200.195	30%
Patient + Impulsive	0.010	229.288	55%	0.010	291.644	0%
Prudent + Impatient	3.666	0.000	100%	0.010	161.606	35%
Prudent + Impulsive	0.500	169.941	55%	0.500	64.506	65%

注) いずれの実験結果も有効回答率は 100%であった。

将来志向的な値が推定された。同様に、Patient + Impulsive ($k = 0.010$) や Prudent + Impatient ($k = 0.010$) でも、各形容詞を単体で与えた場合よりも極端な将来志向の値となった。一方、Prudent + Impulsive では $k = 0.500$ となり、Prudent 単体 ($k = 0.069$) よりも現在志向的な値へと変化した。これらの結果から、Llama3.2 では異なる選好の形容詞を組み合わせた場合に、単体で与えた場合とは異なる相互作用が生じているものの、組み合わせによって予測困難な変化が生じており、形容詞の意味を適切に解釈した上での打ち消し合いとは言い難い。

5.3 人間との比較

人間の推定値の一例 ($k = 2.6 \pm 0.94$) [17] を参考として、本実験の結果と比較する。この被験者は $k = 1.7 \sim 3.5$ 程度の時間選好を示している。表 5 において、形容詞を与えない場合、GPT-4o は $k = 0.010$ 、Llama3.2 は $k = 3.664$ であった。GPT-4o はこの被験者に対し、極めて将来志向的である点で大きく異なる。一方、Llama3.2 はこの被験者の範囲 ($k = 1.7 \sim 3.5$ 程度) に近い値であり、比較的近い特徴を示している。表 5 において、現在志向の形容詞を与えた場合、GPT-4o では $k = 3.666$ となり、この被験者の範囲に近い値へと変化した。Llama3.2 で

は $k = 0.500$ となり、この被験者よりも将来志向的な値であった。これは形容詞を与えない場合 ($k = 3.664$) から低下しており、形容詞の意味とは矛盾する変化である。表 5 において、将来志向の形容詞を与えた場合、GPT-4o では $k = 0.010$ が維持された。元々極めて将来志向的であったため、さらなる変化の余地がなかったと考えられる。Llama3.2 では $k = 0.069 \sim 0.500$ となり、形容詞の意味と一致する将来志向への変化が見られた。ただし、いずれの値もこの被験者の範囲 ($k = 1.7 \sim 3.5$ 程度) からは乖離している。

6 失敗分析

6.1 Llama3.2 の回答の一致度の低さ

Llama3.2 のリスク選好実験では、多くの条件で一致度が 0% となった。これは、選択肢の提示順序を入れ替えると異なる選択肢を選ぶことを意味しており、回答が選択肢の順序に影響されていたことが考えられる。

6.2 Llama3.2 におけるリスク選好パラメータの非変化

Llama3.2 のリスク選好実験では、Adventurous 単独の場合

を除き、ほぼ全ての条件でパラメータが $\alpha \approx 1$, $\beta \approx 1$, $\lambda = 1.000$, $\gamma = 1.000$ のまま変化しなかった。中立な形容詞、リスク回避的な形容詞、リスク愛好的な形容詞 (Audacious 以外)、および形容詞を二つ与えた場合や選好の異なる形容詞を組み合わせた場合のいずれにおいても、パラメータは形容詞を与えない場合の値付近に留まった。Llama3.2 のリスク選好実験では、多くの条件で一致度が 0% となったため、正しく選択肢を評価できていなかった可能性もあるが、Llama3.2 が形容詞に応じてリスク選好を変化させる能力が限定的であることを示している。

6.3 Llama3.2 における形容詞の意味と矛盾した変化

Llama3.2 の時間選好実験では、現在志向の形容詞 (Impatient/Impulsive) を与えた場合に $k = 0.500$ となり、形容詞を与えない場合 ($k = 3.664$) よりもむしろ将来志向的な値へと変化した。これは現在志向の形容詞の意味とは矛盾する変化であり、Llama3.2 が現在志向の形容詞を正しく解釈できていなかった可能性を示している。また、選好の異なる形容詞を組み合わせた場合 (Patient + Impatient, Patient + Impulsive, Prudent + Impatient) でも $k = 0.010$ と極端に将来志向的な値が推定され、GPT-4o で観察されたような選好の打ち消し合いは見られなかった。これらの結果から、Llama3.2 では形容詞の意味を適切に解釈して選好を変化させる能力が限定的であることが考えられる。

6.4 形容詞を与えてもパラメータが変化しなかった場合

GPT-4o のリスク選好実験では、リスク愛好的な形容詞 (Adventurous/Audacious) が λ を 1.000 から 5.000 へ上昇させた一方、リスク回避的な形容詞 (Cautious/Conservative) は $\lambda = 1.000$ を維持した。リスク回避的な形容詞は損失回避を強める方向に作用することが期待されたが、そのような変化は観察されなかった。これは $\lambda \geq 1$ という制約により、形容詞を与えない場合の $\lambda = 1.000$ からさらに低下する余地がなかったためである。同様に、時間選好実験でも、形容詞を与えない場合で $k = 0.010$ と極めて将来志向的であったため、将来志向の形容詞 (Prudent/Patient) によるさらなる変化の余地がなかった。

7 本研究の限界点

本研究にはいくつかの限界点がある。第一に、使用した LLM と形容詞が限定的である点である。本研究では、付録も含めると 8 種類の LLM とリスク選好および時間選好に関してそれぞれ 6 種類の形容詞を使用した。より多様な LLM や形容詞を検討することで、より包括的な知見を得られる可能性がある [9]。第二に、パラメータの推定範囲に関する限界である。本研究では $\lambda \geq 1$ という制約を設けたため、形容詞を与えない場合に $\lambda = 1.000$ であった GPT-4o では、リスク回避的な形容詞による λ の低下を観察できなかった。同様に、時間選好パラメータ k についても、GPT-4o の形容詞を与えない場合が $k = 0.010$ と推定範囲の下限付近であったため、将来志向の形容詞による変化を観察できなかった。これらの問題を回避するためには、

形容詞を与えない場合のパラメータが推定範囲の境界付近に位置しない LLM を選択するか、異なるプロンプト設計によってベースラインのパラメータを調整することが考えられる。

8 おわりに

本研究では、大規模言語モデル (LLM) の内部的な意思決定特性を行動経済学の理論枠組みに基づいて測定し、さらにプロンプト設計を通じて特性を付与・調整する方法を検討した。具体的には、GPT-4o と Llama3.2 を対象に、リスク選好および時間選好を表す形容詞をプロンプトに含めることで、選好がどのように変化するかをパラメータ推定によって分析した。

リスク選好実験では、GPT-4o ではリスク愛好的な形容詞 (Adventurous/Audacious) により λ が 1.000 から 5.000 へ上昇し、損失回避傾向が強化された。リスク回避的な形容詞 (Cautious/Conservative) では $\lambda = 1.000$ が維持されたが、 β が上昇し損失への感応度が高まった。Llama3.2 では一部の条件でのみパラメータが変化し、形容詞に応じて選好を変化させる能力が限定的であった。形容詞を二つ与えた場合、GPT-4o では効果の累積が見られず、選好の異なる形容詞を組み合わせた場合と相互に打ち消し合い中立に近い選好が表現された。

時間選好実験では、GPT-4o では現在志向の形容詞 (Impatient/Impulsive) により k が 0.010 から 3.666 へ上昇し、形容詞の意味と一致する変化が観察された。Llama3.2 では将来志向の形容詞により k が低下したが、現在志向の形容詞でも k が低下し、形容詞の意味と矛盾した変化が見られた。形容詞を二つ与えた場合、GPT-4o では効果の累積が見られなかった一方、Llama3.2 では応答パターンに予測困難な変化が生じた。選好の異なる形容詞を組み合わせた場合、GPT-4o では条件によって異なる形容詞が優勢に働くか、または相互に打ち消し合う結果となった。

本研究ではいくつかの課題も明らかになった。Llama3.2 のリスク選好実験では、多くの条件で一致度が 0% となり、選択肢の提示順序に回答が影響されていたほか、形容詞を与えても選好パラメータがほとんど変化せず、形容詞に応じて選好を変化させる能力が限定的であることが示された。また、Llama3.2 の時間選好実験では、現在志向の形容詞を与えた場合に k が低下するという、形容詞の意味と矛盾した変化が観察された。さらに、GPT-4o では形容詞を与えない場合の λ や k がすでに推定範囲の境界付近にあったため、リスク回避的な形容詞や将来志向の形容詞による変化を十分に観察できなかった。

今後の研究の方向性として、以下の点が挙げられる。第一に、より多様な LLM や形容詞を用いた検証を行うことで、知見の一般化可能性を高めることが考えられる。第二に、形容詞を与えない場合のパラメータが推定範囲の境界付近に位置しない LLM を選択するか、プロンプト設計によってベースラインのパラメータを調整することで、形容詞による変化をより明確に観察できる可能性がある。

本研究の結果は、意思決定特性を表す形容詞をプロンプトに含めることで LLM のリスク選好および時間選好を一定程度制

御できることを示した。この知見は、LLM を意思決定支援や行動シミュレーションに応用する際に、モデルの選好特性を意図的に調整するための一手法として活用できる可能性がある。

謝 辞

本研究の一部は、JSPS 科研費 24K03046, 24K23877 の助成を受けたものです。ここに記して謝意を示します。

文 献

- [1] Chun-Wei Chiang, Zhuoran Lu, Zhuoyan Li, and Ming Yin. Enhancing ai-assisted group decision making through llm-powered devil's advocate. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*, pp. 103–119, 2024.
- [2] Ayato Kitadai, Sinndy Dayana Rico Lugo, Yudai Tsurusaki, Yusuke Fukasawa, and Nariaki Nishino. Can ai with high reasoning ability replicate human-like decision making in economic experiments? *Group Decision and Negotiation*, pp. 1–24, 2025.
- [3] 馬田光琉, 山本修平. 行動経済学的特性を付与した大規模言語モデルによる消費行動予測. 人工知能学会全国大会論文集 第 39 回 (2025), pp. 1D3OS24a03–1D3OS24a03. 一般社団法人 人工知能学会, 2025.
- [4] Jingru Jessica Jia, Zehua Yuan, Junhao Pan, Paul McNamara, and Deming Chen. Decision-making behavior evaluation framework for llms under uncertain context. *Advances in Neural Information Processing Systems*, Vol. 37, pp. 113360–113382, 2024.
- [5] 筒井義郎, 佐々木俊一郎, 山根承子, グレグ・マルデフ. 行動経済学入門. 東洋経済新報社, 2017.
- [6] Charles A Holt and Susan K Laury. Risk aversion and incentive effects. *American economic review*, Vol. 92, No. 5, pp. 1644–1655, 2002.
- [7] Kris N Kirby and Nino N Maraković. Delay-discounting probabilistic rewards: Rates decrease as amounts increase. *Psychonomic bulletin & review*, Vol. 3, No. 1, pp. 100–104, 1996.
- [8] 吉川克正, 大萩雅也, 高山隼矢. 大規模言語モデルの非対称的意思決定特性：プロスペクト理論要素の実証分析. 言語処理学会 第 31 回年次大会 発表論文集, pp. 195–200, 2025.
- [9] 吉川克正, 大萩雅也, 高山隼矢. 大規模言語モデルは人間の時間選好性を再現できるか? 人工知能学会全国大会論文集 第 39 回 (2025), pp. 2A1GS1001–2A1GS1001. 一般社団法人 人工知能学会, 2025.
- [10] Yuta Imasaka and Hideo Joho. Effect of llm's personality traits on query generation. In *Proceedings of the 2024 Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region*, pp. 249–258, 2024.
- [11] Yuxuan Li, Hirokazu Shirado, and Sauvik Das. Actions speak louder than words: Agent decisions reveal implicit biases in language models. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*, pp. 3303–3325, 2025.
- [12] Jessica Maria Echterhoff, Yao Liu, Abeer Alessa, Julian McAuley, and Zexue He. Cognitive bias in decision-making with llms. In *Findings of the association for computational linguistics: EMNLP 2024*, pp. 12640–12653, 2024.
- [13] 角田康明, 竹内孝, 鹿島久嗣. 大規模言語モデルにおける認知バイアスの軽減. 人工知能学会全国大会論文集 第 38 回 (2024), pp. 2T5OS5b02–2T5OS5b02. 一般社団法人 人工知能学会, 2024.
- [14] Amos Tversky and Daniel Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, Vol. 5, No. 4, pp. 297–323, 1992.
- [15] 川越敏司. 「意思決定」の科学：なぜ、それを選ぶのか. 講談社, 2020.
- [16] Adam S Booij, Bernard MS Van Praag, and Gijs Van De Kuilen. A parametric analysis of prospect theory's functionals for the general population. *Theory and Decision*, Vol. 68, No. 1, pp. 115–148, 2010.
- [17] Nicolas Schweighofer, Kazuhiro Shishida, Cheol E Han, Yasumasa Okamoto, Saori C Tanaka, Shigeto Yamawaki, and Kenji Doya. Humans can adopt optimal discounting strategy under real-time constraints. *PLoS computational biology*, Vol. 2, No. 11, p. e152, 2006.