

日本語因果条件文における Jeffrey table の妥当性と 確率的性質の検証

Validity and Probabilistic Properties of Jeffrey Table in Japanese Causal Conditionals

吉沢 栄貴[†], 高橋 達二[‡]

Hideki Yoshizawa, Tatsuji Takahashi

[†] 東京電機大学大学院, [‡] 東京電機大学 理工学部

Graduate School of Tokyo Denki University, School of Science and Engineering, Tokyo Denki University
tatsujit@mail.dendai.ac.jp

概要

「 p ならば q 」という形式をとる条件文を人間が解釈する際、論理学上の古典的な定義と人間の直感的解釈に相違があることが知られており、加えて近年新たなアプローチの真理値表として Jeffrey table が提唱されている。また条件文の確率的判断についての先行研究では Over らによる研究がある。この研究での実験を参考に日本語条件文で実験を行い、日本語での条件文の確率的解釈の分析および、Jeffrey table の妥当性を検証を行なった。

キーワード：条件文, 主観的確率, Jeffrey table

1. 序論

「 p ならば q 」といった形式をもつ条件文は、高次認知においては推論や判断、意思決定で中心的な役割を果たし、また言語においては社会的ルールや因果関係、論理的関係の表現において用いられる [1]。AI スピーカーなどの普及により自然言語での機械への指示出しの機会が増えてきているが、現状では条件文を用いてルールを学習させるなどではできないため、人間の条件文の適切なモデリングはより柔軟なパーソナルアシスタントの実装などにおいて社会的に大きな意味があると考えられる。

条件文については、古典的二値論理では前件 p が真 (True)、後件 q が偽 (False) の場合は偽、それ以外の場合では真であるとする実質含意 (material implication) として定義されている。しかしこの論理は人間の直感的解釈には相違があることはこれまでの研究でも示唆されている [1]。これのほかに、前件 p が偽であるときの真理値に、新たな値不定 (Uncertain) をとるとする de Finetti が提唱した de Finetti table や、同じく前件 p が偽のときは、前件 p と後件 q についての条件付き確率 $P(q|p)$ をとるとする Jeffrey が提唱した Jeffrey table といったアプローチも存在している [2, 3]。de

Finetti table や条件付き確率はこれまでも多くの論文で触れられていることがあるが [4, 5]、Jeffrey table については紹介されることが少なく、それに伴い発表されている論文も少ないのが現状である [6]。しかし、条件文をモデリングする際には de Finetti table のような Uncertain で表すのではなく、確率値のような連続値で表すべきであることから研究することは有意義である。人間の条件文の解釈を確率的に分析を試みた研究として、Over らの実験 [7] があり、この結果では条件文の確率的解釈として条件付き確率がよくフィットするという結果が得られている。

そこで本研究では、Over らが行なった実験を参考として、条件文「 p ならば q 」の前件 p と後件 q の生起確率高低によって解釈の違いについて変化があるのかを実験により明らかにすることを目的とする。また、確率値として回答を集計することにより Jeffrey table の妥当性についての検証も目的とする。

2. 条件文の解釈についての変化

論理学において真理値は「真」、「偽」の2つで構成される二値論理が主に用いられている。この二値論理を使い、論理学では条件文は実質含意 (material implication) として既定されており、これは $\neg p \vee q$ と等しい。これは前件 p が真、後件 q が偽のとき以外の場合にはすべて条件文が真、すなわち前件 p が偽であるときは後件 q がどちらの真理値でも条件文は真であるとしているものである。しかし、この解釈を人間に当てはめようとするとき直感との解釈の相違が生じてしまうと考えられており、二値論理では人間の直感的な論理を表現しようとする際に正確さに欠く部分がある。そのため、真偽について分からない場合についての不確実性を考慮した論理体系が必要であるとされている。

そこで「真」、「偽」の他に第3の真理値「不定」を加えた真理値表、「欠陥真理値表」という概念が de

表1 代表的な解釈の真理値表

p	q	実質含意	条件付き事象	双条件付き事象
T	T	T	T	T
T	F	F	F	F
F	T	T	U	F
F	F	T	U	U

Finetti により提唱され、これは新パラダイム推論心理学により妥当であるとされており [8, 9], 別名 de Finetti table とも呼ばれることがある。この「不定」を用いた条件付き事象 (conditional event) という解釈を導入することで相違の解消が提唱されている。この条件付き事象とは、実質含意とは異なり前件 p が偽であるときは後件 q の真偽に関わらず不定とするものである。この解釈は条件文の形式に対応しているものと考えられている。表1に代表的な解釈の真理値表をまとめる。

3. Jeffrey table

Jeffery table は、de Finetti table の影響を受け Jeffery によって提唱された真理値表である [3]。de Finetti table は前件 p が偽であるとき後件 q の真偽に関わらず不定とするものである。Jeffery table は前件 p が偽であるとき後件 q が真でも偽でも真理値として $P(q | p)$ の値をとると定義している。以下の表2に jeffery table を示す。

表2 Jeffrey table

p	q	Value
T	T	1
T	F	0
F	T	$P(q p)$
F	F	$P(q p)$

4. 条件文における確率モデル

条件文の解釈について確率的にモデリングを行う際、条件文の前件 p , 後件 q の真偽の組み合わせについて計算可能であることがこれまでの研究により示されている。以下に代表的な条件文の確率モデルを示す。

表3 代表的な確率モデル

モデル	確率式
Conjunctive probability:	$P(p \& q) = P(TT)$
Material conditional:	$P(MC) = P(TT) + P(FT) + P(FF)$
Conditional probability:	$P(q p) = P(TT) / [P(TT) + P(TF)]$
Delta-p rule:	$\Delta p = P(q p) - P(q \neg p)$

5. 実験について

5.1 Over らによる先行研究

本研究の先行研究として、Over らが2007年に行った実験が挙げられる [7]。この研究では、条件文の前件 p , 後件 q の生起確率の高低によりタイプを4種 (HH, HL, LH, LL) に分けた上で使用し実験を行っていた。この実験においては以下の3つのタスクを行い、確率モデルについての検証を行っていた。

1. 条件文の前件 p , 後件 q の真偽の組み合わせについて4つのケースにおける確率を合計で100になように回答する結合分布タスク
2. 条件文が確率的にどれだけ正しいと言えるかを回答する確率判断タスク
3. 条件文の前件 p , 後件 q の間にある因果関係の強さを回答する因果強度タスク

5.2 本研究での実験内容

本研究で行ったプレテスト及び本実験はどちらも実験参加者をクラウドソーシングで募集し、オンライン調査ツール Qualtrics 上で実験を行なった。これは実験参加者・主催者双方における負担を軽減しつつ分析に当たり十分なデータ数を得ることを目的とするものである。

5.3 プレテスト

以降で述べる本実験のために、条件文の前件32種と後件32種の全64文の確率を測定した。ここで用意した文は Over らが使用した条件文を参考にし [7], 「10年以内に起こりそうなもの」という前提で作られたものを「日本で10年以内に起こりそうなもの」となるように設定した。

5.3.1 実験手順

実験参加者は100名募集し、提示した文の内容が今後10年以内に日本で起こる確率を0~100の範囲の数値をとるスライダーで回答する形式を採用した。また、文はランダム順で提示した上で全64文について行った。

5.3.2 実験結果

実験参加者の総数は、研究に使用することに同意しない2名を除いた98名(平均年齢38.37, 標準偏差10.79)となった。この結果により、日本語における各文の確率を得ることができ、それらが使用されている条件文のタイプ別分類を行うことが可能となった。この結果を基に付録A章に示す条件文16種を選択した。

5.4 本実験手順

本実験ではクラウドソーシングにより実験参加者を募集し、オンライン調査ツール Qualtrics 上で実験に回答してもらう方式を採用した。本実験では Over による3つのタスクに加え、Jeffrey table タスクを加えた4つのタスクを行った。全てのタスクにおいて、全16種の条件文をランダムに表示することで繰り返し回答を収集した。この実験で使用した条件文は付録Aの章に示すものである。各タスクの手順については以下に示す。

- 結合分布タスクでは、16の条件文について、TT, TF, FT, FFの確率を回答する。各ケースの確率をスライダーで0から100で回答し、4つのケースの確率の合計が100となるように回答する。
- 確率判断タスクでは、16種の条件文について条件文が正しいと思われる確率を回答する。スライダーで0から100の間で回答する。
- 因果強度タスクでは、16種の条件文について条件文の前件 p と後件 q の間にある因果関係の強さをスライダーで1から5の5段階で回答する。
- Jeffrey table タスクでは、16種の条件文について、TT, TF, FT, FFの確率を回答する。これは結合分布タスクとは異なり4つのケースでの合計を100にする必要はなく、各ケースにおいて0から100の間の確率値をスライダーにより回答する。

また、上記の4つのタスクは結合分布タスクを行うグループ、確率判断タスクと因果強度タスクを行うグループ、Jeffrey table タスクを行うグループをカウンターバランスをとった上で割り振りを行った。

6. 実験結果

本実験では合計で300人の実験参加者を募集し、100人ずつ3つのグループを割り振った。また、実験内には実験参加者が質問文を読んでいるかを確認するIMCを埋め込んだ。IMCは質問にある定まった値を回答するように指定することで実現した。このIMCで正しく回答していない回答、及び回答を研究に使用するこ

とを希望しない回答を除いた結果、結合分布タスクを行うグループでは72人、確率判断タスクと因果強度タスクを行うグループでは98人、Jeffrey table タスクを行うグループでは93人の有効データを得ることが出来た。結合分布タスクにおける結果を表4、確率判断タスクにおける結果を表5、因果強度タスクにおける結果を表6、Jeffrey table タスクにおける結果を表7に示す。

表4 結合分布での各真理値の平均(SD)

TT	q	
	q -High	q -Low
p -High	36.48(13.1)	19.14(5.0)
p -Low	17.60(5.2)	10.72(3.8)
TF	q	
	q -High	q -Low
p -High	31.88(13.9)	48.49(12.1)
p -Low	10.16(3.1)	16.33(6.3)
FT	q	
	q -High	q -Low
p -High	15.20(6.0)	9.18(3.7)
p -Low	45.15(16.2)	26.00(6.0)
FF	q	
	q -High	q -Low
p -High	16.44(4.6)	23.19(7.1)
p -Low	27.10(9.2)	46.96(4.5)

表5 確率判断タスクでの平均(SD)

p	q	
	High	Low
High	55.02(6.0)	32.47(9.0)
Low	55.29(15.6)	40.10(11.9)

表6 因果強度タスクでの平均(SD)

p	q	
	High	Low
High	2.90(0.3)	2.02(0.4)
Low	2.95(0.7)	2.31(0.4)

表7 Jeffrey table タスクでの平均(SD)

TT	q	
	q -High	q -Low
p -High	69.10(5.6)	55.70(7.2)
p -Low	68.76(9.1)	57.71(8.5)
TF	q	
	q -High	q -Low
p -High	39.29(9.9)	49.55(2.3)
p -Low	30.46(2.9)	39.80(8.8)
FT	q	
	q -High	q -Low
p -High	33.86(7.7)	29.46(3.0)
p -Low	46.95(5.7)	34.69(5.5)
FF	q	
	q -High	q -Low
p -High	56.69(7.8)	61.78(3.2)
p -Low	48.01(4.4)	61.98(6.1)

7. 考察

7.1 先行研究との比較

5.1 節において先行研究として Over らが 2007 年に行なった実験を挙げた [7]. Over らは実験で集計した回答を基に, 分散分析, 相関分析, 重回帰分析を行っていた. 本項ではこの分析手法に倣い, 同様の分析を行い非日本語圏で行われた実験との比較を試みる.

7.1.1 分散分析による分析

Over らは分散分析を行う際, 前件 p の生起確率の高低, 後件 q の生起確率の高低についての二要因分散分析を 2 種類行っていた. その 2 種類は, 実験参加者に関する分析を F_1 として, 条件文に関する分析を F_2 として行なうものであった.

Over らの実験では, 結合分布タスクにおいて F_1 では TF の前件及び交互作用について, F_2 では TF の前件, TT, TF, FT の交互作用について $\alpha = 0.05$ 水準で有意差が認められず, それら以外では有意差が認められる結果だった. 確率判断タスクでは, F_1 では前件, 後件, 交互作用の全てで有意差が認められたが, F_2 では全てで認められない結果だった. この結果は因果強度タスクでも同様であった.

本研究では Over らに倣い, 各タスクについての分散分析を行い有意差を測定した. 表 8 から表 15 に各タスクでの F_1 , F_2 の結果をまとめたものを示す.

本研究の結果では, F_1 での分析では結合分布タスクでの全ての真理値, 確率判断タスク及び因果強度タスクにおいて全てで前件, 後件, 相互作用について有意差があると認められた. F_2 での分析では, 結合分布タスクでは全ての真理値の相互作用, 確率分布タスクでは前件及び相互作用, 因果強度タスクにおいても前件及び相互作用で有意差が認められなかった. 特に Over らの分析では見られなかった確率判断タスクと因果強度タスクでの F_2 の後件での有意差が本研究では認められたことから, 日本語条件文と非日本語条件文において後件の生起確率による解釈の違いが生まれているのではないかと考えられる.

表 8 結合分布タスクでの F_1 の結果 ($\alpha = 0.05$)

	F 値	p 値
TT		
前件	110.2	$p < .001$
後件	105.2	$p < .001$
交互作用	24.08	$p < .001$
TF		
前件	287.5	$p < .001$
後件	91.37	$p < .001$
交互作用	22.94	$p < .001$
FT		
前件	394	$p < .001$
後件	93.34	$p < .001$
交互作用	26.79	$p < .001$
FF		
前件	128.6	$p < .001$
後件	97.12	$p < .001$
交互作用	25.8	$p < .001$

表 9 結合分布タスクでの F_2 の結果 ($\alpha = 0.05$)

	F 値	p 値
TT		
前件	12.480	0.00413
後件	9.815	0.00865
交互作用	1.828	0.20126
TF		
前件	29.878	$p < .001$
後件	5.337	0.039462
交互作用	1.122	0.310399
FT		
前件	25.070	0.000306
後件	7.267	0.019462
交互作用	1.975	0.185231
FF		
前件	26.866	$p < .001$
後件	16.050	0.001742
交互作用	3.897	0.071842

表 10 確率判断タスクでの F_1 の結果 ($\alpha = 0.01$)

	F 値	p 値
前件	14.01	$p < .001$
後件	236.5	$p < .001$
交互作用	8.256	0.00499

表 11 確率判断タスクでの F_2 の結果 ($\alpha = 0.05$)

	F 値	p 値
前件	0.499	0.4936
後件	11.336	0.0056
交互作用	0.432	0.5236

表 12 因果強度タスクでの F_1 の結果 ($\alpha = 0.05$)

	F 値	p 値
前件	9.922	0.00217
後件	128.1	$p < .001$
交互作用	4.307	0.0406

表 13 因果強度タスクでの F_2 の結果 ($\alpha = 0.05$)

	F 値	p 値
前件	0.487	0.49836
後件	9.805	0.00867
交互作用	0.258	0.62081

表 14 Jeffrey table タスクでの F_1 の結果 ($\alpha = 0.05$)

	F 値	p 値
TT		
前件	0.591	0.444
後件	60.8	$p < .001$
交互作用	1.296	0.258
TF		
前件	44.53	$p < .001$
後件	56.34	$p < .001$
交互作用	0.151	0.698
FT		
前件	69.58	$p < .001$
後件	34.00	$p < .001$
交互作用	10.98	0.00132
FF		
前件	10.69	0.00152
後件	41.04	$p < .001$
交互作用	10.69	0.00152

表 15 Jeffrey table タスクでの F_2 の結果 ($\alpha = 0.05$)

	F 値	p 値
TT		
前件	0.047	0.83270
後件	10.062	0.00804
交互作用	0.093	0.76598
TF		
前件	7.304	0.0192
後件	8.129	0.0146
交互作用	0.018	0.8946
FT		
前件	10.108	0.00793
後件	8.358	0.01355
交互作用	1.855	0.19821
FF		
前件	2.242	0.16017
後件	11.325	0.00562
交互作用	2.463	0.14253

7.1.2 相関分析

Over らの先行研究では諸確率モデルと確率分布タスク及び因果強度判断タスクの回答で相関を取っており、その結果を表 16 に示す。

本研究の結果は表 17 に確率判断タスク及び因果強度タスクの相関を示す。

Over らの結果では確率判断タスク、因果強度タスク双方において最も相関が高かったのは条件付き確率 $P(q|p)$ であり、本研究の結果においても同様の結果が得られたが確率分布タスクにおいては有意差が認められなかった。これは本研究の実験設定上の問題である、タスク間においての参加者の対応が取れておらず回答とモデル値を問題ごとに平均をとり相関分析を行ったことによるものであると考えられる。

表 16 Over ら (2007) での相関 (対応あり)

	確率判断	因果強度
代表的モデル		
$P(p\&q)$	0.88*	0.86*
$P(MC)$	0.77*	0.71*
$P(q p)$	0.91*	0.87*
Δp	0.72*	0.73*
その他モデル		
$P(p)$	0.46*	0.49*
$P(q)$	0.77*	0.73*
$P(q \neg p)$	0.11	0.07

* $p < .05$

表 17 本研究での相関 (対応なし)

	確率判断	因果強度
代表的モデル		
$P(p\&q)$	0.43*	0.41
$P(MC)$	0.51*	0.48
$P(q p)$	0.65	0.62*
Δp	0.49	0.50*
その他モデル		
$P(q p)$	0.46	0.43
$P(p)$	-0.18	-0.16
$P(q)$	0.62*	0.60*
$P(q \neg p)$	0.45	0.42

* $p < .05$

7.1.3 重回帰分析

Over らの先行研究では確率モデルのうち $P(p), P(q|p), P(q|\neg p)$ を説明変数、確率分布タスク及び因果強度判断タスクの回答を目的変数として重回帰分析を行い、確率モデルの偏回帰係数を測定していた。この結果は表 18 に示す。

本研究での結果は表 19 に重回帰分析の結果を示す。この分析において条件付き確率 $P(q|p)$ が最も結果となる偏回帰係数が高い結果となり、相関分析のときと同様に Over らの結果と同様となることが確認できた。

また、有意差についても有意水準 $\alpha = 0.05$ で認められる結果となり Over らの結果と一致した。

しかし、この分析においても確率モデル値は質問ごとの平均をとったものを使用、つまり全 16 種の値を使用してしまっているため分析の結果として有効であるとは言い難いと考えられる。

表 18 Over ら (2007) での重回帰分析 (対応あり)

確率モデル	確率判断	因果強度
$P(p)$	0.14*	0.19*
$P(q p)$	0.93*	0.88*
$P(q \neg p)$	-0.20*	-0.23*

* $p < .05$

表 19 重回帰分析の結果 (対応なし)

確率モデル	確率判断	因果強度
$P(p)$	-0.09*	-0.35*
$P(q p)$	0.66*	2.68*
$P(q \neg p)$	-0.06	-0.29

* $p < .05$

7.2 確率分布間の距離測定

今回の実験では、3 つのタスクグループに実験参加者を割り振り実験を行なった。各タスクにおいて十分な実験参加者を確保することを目的にしたことであるが、これは実験にかかる時間が長引くことからの参加者の集中力の低下による途中離脱を防止するという側面も含む。

そのため Over らの実験とは異なり、各参加者のタスク間の回答の対応がつかない結果となってしまったため、一概に比較及び検討をすることは難しいと考えられた。そこで、確率分布間の距離の差を測ることが可能な Kullback-Leibler divergence を用いることで確率モデルの妥当性を検証することとした。

この分析では結合分布タスクの結果から計算した確率モデル値と、今回の実験において確率値として回答を集計した確率分布タスク及び、Jeffrey table タスクの FT, FF の結果との比較することを行い、各回答におけるモデルの当てはまりの良さの計測し、Jeffrey table の妥当性について検証を試みた。各タスクの回答の確率モデル値間の距離及び、全 16 問で平均した結果を表 20 から 25 に示す。

HH から LL までの全タイプを平均した結果について着目すると、確率判断タスクと Jeffrey table タスク

の FT の回答については距離に近いモデルは $P(q|\neg q)$ となり、Jeffrey table タスクの FF の回答についての距離に近いモデルは $P(q|p)$ となった。Jeffrey table の FF については元来の定義の通りであるため望ましい結果であると言えるが、その反面、FT については定義とは異なる結果となった。 $P(q|\neg q)$ は言葉で表すと、 p が起こらないときに q が起きる確率 ということであるため、FT の意味という観点からの妥当性は認められると考えられるが確証は得られていないため、それに関する検証の必要が求められる。

表 20 確率判断タスクの回答分布と諸モデル間の距離

HH	$P(q p)$	$P(q \neg p)$	Δp	$P(p\&q)$	$P(q p)$	$P(MC)$	$P(p)$	$P(q)$
Q1	0.17	0.08	0.40	0.21	0.11	0.40	0.04	0.12
Q2	0.19	0.04	0.50	0.15	0.07	0.46	0.02	0.14
Q3	0.48	0.07	0.41	0.03	0.02	0.88	0.14	0.41
Q4	0.28	0.04	0.41	0.10	0.05	0.59	0.06	0.24
Mean	0.28	0.06	0.43	0.12	0.06	0.58	0.07	0.23
HL	$P(q p)$	$P(q \neg p)$	Δp	$P(p\&q)$	$P(q p)$	$P(MC)$	$P(p)$	$P(q)$
Q1	0.06	0.11	1.94	0.17	0.13	0.10	0.82	0.07
Q2	0.16	0.26	1.67	0.33	0.26	0.04	0.56	0.16
Q3	0.03	0.08	2.33	0.08	0.04	0.23	0.97	0.03
Q4	0.24	0.34	1.58	0.41	0.34	0.06	0.52	0.24
Mean	0.12	0.20	1.88	0.25	0.19	0.11	0.72	0.12
LH	$P(q p)$	$P(q \neg p)$	Δp	$P(p\&q)$	$P(q p)$	$P(MC)$	$P(p)$	$P(q)$
Q1	0.11	0.06	0.39	1.16	1.02	0.24	0.68	0.03
Q2	0.12	0.33	1.04	0.59	0.37	0.65	0.18	0.18
Q3	0.18	0.03	0.19	1.57	1.38	0.18	0.96	0.06
Q4	0.11	0.03	0.45	1.20	0.97	0.29	0.61	0.01
Mean	0.13	0.11	0.52	1.13	0.94	0.34	0.61	0.07
LL	$P(q p)$	$P(q \neg p)$	Δp	$P(p\&q)$	$P(q p)$	$P(MC)$	$P(p)$	$P(q)$
Q1	0.02	0.06	0.29	0.30	0.35	1.77	0.08	0.02
Q2	0.05	0.06	0.23	0.27	0.30	1.99	0.04	0.02
Q3	0.13	0.03	0.08	0.04	0.06	2.72	0.05	0.10
Q4	0.13	0.28	0.53	0.57	0.74	0.99	0.40	0.15
Mean	0.08	0.11	0.28	0.29	0.36	1.87	0.14	0.07

表 21 確率判断タスクの回答分布と諸モデル間の距離の平均

$P(q p)$	$P(q \neg p)$	Δp	$P(p\&q)$	$P(q p)$	$P(MC)$	$P(p)$	$P(q)$
0.15	0.12	0.78	0.45	0.39	0.72	0.38	0.12

表 22 Jeffrey タスク (FT) の回答分布と諸モデル間の最短距離

HH	$P(q p)$	$P(q \neg p)$	Δp	$P(p\&q)$	$P(q p)$	$P(MC)$	$P(p)$	$P(q)$
Q1	0.71	0.14	0.68	0.03	0.09	1.25	0.29	0.69
Q2	0.60	0.09	0.63	0.01	0.05	1.09	0.21	0.58
Q3	0.73	0.13	0.83	0.04	0.10	1.28	0.29	0.72
Q4	1.05	0.30	1.20	0.17	0.29	1.70	0.53	1.09
Mean	0.78	0.16	0.84	0.06	0.13	1.33	0.33	0.77
HL	$P(q p)$	$P(q \neg p)$	Δp	$P(p\&q)$	$P(q p)$	$P(MC)$	$P(p)$	$P(q)$
Q1	0.09	0.16	1.86	0.23	0.16	0.07	0.71	0.09
Q2	0.03	0.05	2.23	0.08	0.04	0.19	0.97	0.03
Q3	0.07	0.16	1.93	0.23	0.14	0.10	0.72	0.06
Q4	0.03	0.08	2.37	0.07	0.03	0.27	1.02	0.02
Mean	0.05	0.11	2.10	0.15	0.09	0.16	0.85	0.05
LH	$P(q p)$	$P(q \neg p)$	Δp	$P(p\&q)$	$P(q p)$	$P(MC)$	$P(p)$	$P(q)$
Q1	0.06	0.08	0.46	1.06	0.85	0.33	0.53	0.02
Q2	0.08	0.15	0.74	0.87	0.62	0.46	0.34	0.06
Q3	0.14	0.34	1.04	0.62	0.39	0.69	0.20	0.19
Q4	0.11	0.17	0.76	0.85	0.60	0.52	0.34	0.08
Mean	0.10	0.19	0.75	0.85	0.62	0.50	0.35	0.09
LL	$P(q p)$	$P(q \neg p)$	Δp	$P(p\&q)$	$P(q p)$	$P(MC)$	$P(p)$	$P(q)$
Q1	0.09	0.02	0.19	0.10	0.12	2.59	0.07	0.10
Q2	0.03	0.01	0.18	0.15	0.19	2.18	0.04	0.02
Q3	0.02	0.04	0.28	0.26	0.30	1.88	0.11	0.02
Q4	0.05	0.04	0.22	0.22	0.24	2.11	0.03	0.04
Mean	0.05	0.03	0.22	0.18	0.21	2.19	0.06	0.04

表 23 Jeffrey タスク (FT) の回答分布と諸モデル間距離の平均

$P(q p)$	$P(q \neg p)$	Δp	$P(p\&q)$	$P(q p)$	$P(MC)$	$P(p)$	$P(q)$
0.24	0.12	0.98	0.31	0.26	1.05	0.40	0.24

表 24 Jeffrey タスク (FF) の回答分布と諸モデル間の最短距離

HH	$P(q p)$	$P(q \neg p)$	Δp	$P(p\&q)$	$P(q p)$	$P(MC)$	$P(p)$	$P(q)$
Q1	0.30	0.01	0.56	0.06	0.03	0.66	0.05	0.28
Q2	0.36	0.01	0.57	0.03	0.02	0.77	0.08	0.34
Q3	0.06	0.07	0.67	0.27	0.15	0.29	0.02	0.07
Q4	0.06	0.08	0.62	0.28	0.16	0.26	0.01	0.05
Mean	0.19	0.04	0.61	0.16	0.09	0.49	0.04	0.19
HL	$P(q p)$	$P(q \neg p)$	Δp	$P(p\&q)$	$P(q p)$	$P(MC)$	$P(p)$	$P(q)$
Q1	1.13	1.21	0.71	1.51	1.41	0.33	0.10	1.15
Q2	1.05	1.14	0.80	1.38	1.32	0.32	0.15	1.08
Q3	0.74	0.83	0.89	1.10	0.99	0.13	0.19	0.75
Q4	0.95	1.03	0.80	1.33	1.23	0.24	0.23	0.98
Mean	0.97	1.05	0.80	1.33	1.24	0.26	0.17	0.99
LH	$P(q p)$	$P(q \neg p)$	Δp	$P(p\&q)$	$P(q p)$	$P(MC)$	$P(p)$	$P(q)$
Q1	0.05	0.14	0.66	0.81	0.62	0.37	0.36	0.05
Q2	0.07	0.25	0.87	0.67	0.45	0.54	0.23	0.12
Q3	0.05	0.11	0.47	1.00	0.84	0.29	0.53	0.04
Q4	0.07	0.18	0.61	0.93	0.66	0.51	0.39	0.08
Mean	0.06	0.17	0.65	0.85	0.64	0.43	0.38	0.07
LL	$P(q p)$	$P(q \neg p)$	Δp	$P(p\&q)$	$P(q p)$	$P(MC)$	$P(p)$	$P(q)$
Q1	0.48	0.71	1.04	1.06	1.33	0.41	0.98	0.55
Q2	0.36	0.55	0.84	0.83	1.11	0.57	0.81	0.42
Q3	0.10	0.23	0.56	0.55	0.69	1.04	0.38	0.14
Q4	0.31	0.49	0.77	0.80	1.02	0.63	0.70	0.36
Mean	0.31	0.49	0.80	0.81	1.04	0.66	0.72	0.37

表 25 Jeffrey タスク (FF) の回答分布と諸モデル間距離の平均

$P(q p)$	$P(q \neg p)$	Δp	$P(p\&q)$	$P(q p)$	$P(MC)$	$P(p)$	$P(q)$
0.38	0.44	0.72	0.79	0.75	0.46	0.33	0.40

8. 結論

本研究では、Over らの実験を参考にし日本語条件文の確率的判断の分析及び、確率モデルとの比較による Jeffrey table の妥当性の検証を行なった。Over らの分析手法と同様の分析を行なった結果、概ね Over らの結果と同様条件付き確率 $P(q|p)$ のモデルの結果が良いという結果が得られた。しかし実験設計における参加者の対応について、Over らの実験と相違が生まれてしまったため、一概に同じであると結論づけることは難しく、また Kullback-Leibler divergence による分析の結果では異なるモデルがフィットしていることが明らかになった。そのため、先の問題点を解決及びより詳細な分析を行うことが今後の課題及び展望として挙げられる。

文献

- [1] Mike Oaksford, & Nick Chater, (2010) "Cognition and Conditionals: Probability and Logic in Human Thinking.", *Oxford University Press*.
- [2] David Over, Jean Baratgin, (2016) "The "Defective" Truth Table: Its Past, Present, and Future", *The Thinking Mind*, pp. 15-28.
- [3] Richard Jeffreys, Dorothy Edgington, (1991), "Matter-of-Fact Conditionals", *Proceedings of the Aristotelian Society*, 29, 161-183+185-209.
- [4] Jean Baratgin, David Over, Guy Politzer, (2013) "Uncertainty and the de Finetti tables.", *Thinking & Reasoning* 19, 308-328.
- [5] Jean Baratgin, Guy Politzer, David E. Over, & Tatsuji Takahashi, (2018) "The psychology of uncertainty and three-valued truth tables", *Frontiers in Psychology* 9, 1479.
- [6] Moyun Wang, Mingyi Zhu, (2019) "Evidence for the Jeffrey Table Credibility Ratings for Conditionals Given False Antecedent Cases", *Experimental Psychology*.doi:10.1027/1618-3169/a000443
- [7] David E. Over, Constantinos Hadjichristidis, Jonathan St.B.T. Evans, Simon J. Handley, & Steven A. Sloman, (2007) "The probability of causal conditionals", *Cognitive Psychology* 54, pp. 62-97.
- [8] Politzer, G., Over, D.E., & Baratgin, J. (2010) "Betting on conditionals." *Thinking and Reasoning*, 16, 3, 172-197.
- [9] Over, D.E. (2009) "New Paradigm psychology of reasoning" *Thinking & Reasoning*, 15(4), 431-438.

9. 付録 A

表 26 本研究で使用した条件文 16 種

Type	ID	Antecedent	P(Antecedent)	Consequent	P(Consequent)
HH	HH1	日本で完全自動運転の車が発売される	0.65	マニュアル車の販売台数が少なくなる	0.82
HH	HH2	不妊治療の技術が向上する	0.74	世界の人口が増加する	0.65
HH	HH3	ガソリンの値段が上がる	0.70	交通渋滞が減少する	0.52
HH	HH4	後期高齢者の医療費の窓口負担率が高くなる	0.75	病院での待ち時間が減る	0.53
HL	HL1	自民党の総裁が変わる	0.77	日経平均株価が4万円を上回る	0.25
HL	HL2	地球温暖化が進行する	0.75	東京が水没する	0.26
HL	HL3	より多くの人々が日光浴の時に日焼け止めを使う	0.69	インドア派の人が減少する	0.31
HL	HL4	暴力的ゲームの発売が規制される	0.60	暴力犯罪が減少する	0.35
LH	LH1	喫煙者が増加する	0.19	たばこ税が増税される	0.85
LH	LH2	日本の出生率が増加する	0.29	(次の)選挙で自民党が勝つ	0.75
LH	LH3	車の所有者が増加する	0.36	大気汚染が悪化する	0.70
LH	LH4	日本で死刑が廃止される	0.25	凶悪犯罪率が増加する	0.62
LL	LL1	アサド大統領が暗殺される	0.40	シリアが民主主義の国になる	0.32
LL	LL2	参議院選挙で立憲民主党が過半数を獲得する	0.28	日本がTPPを離脱する	0.39
LL	LL3	離婚手続きが難しくなる	0.30	婚姻率が増加する	0.34
LL	LL4	日本共産党が衆議院の過半数を獲得する	0.15	憲法改正が阻止される	0.49

10. 付録 B

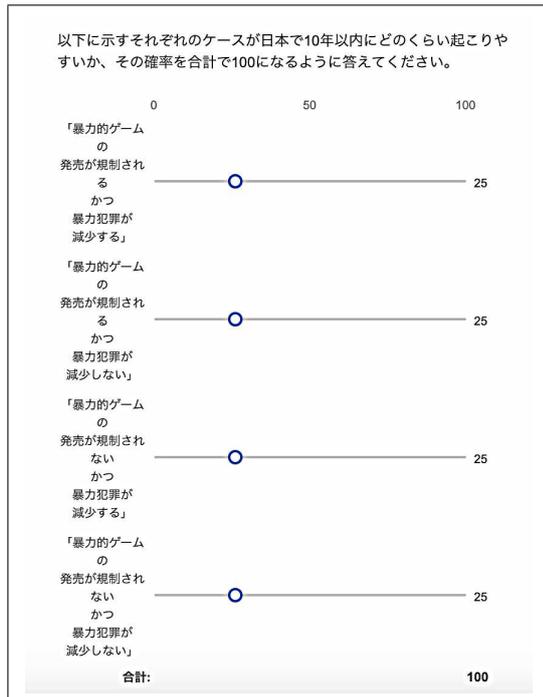


図 1 結合分布タスクの実験画面サンプル

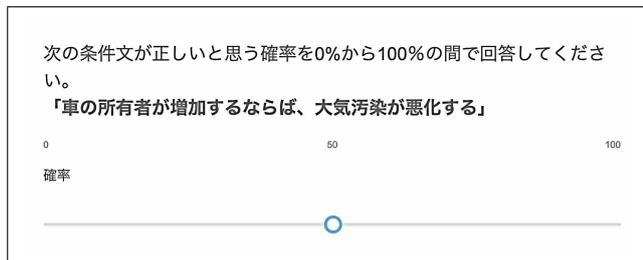


図 2 確率判断タスクの実験画面サンプル

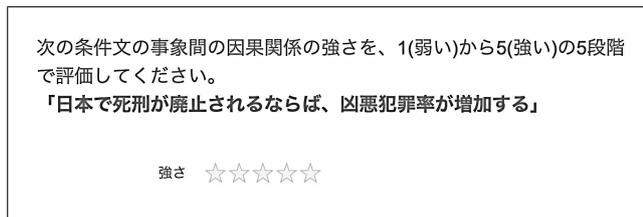


図 3 因果強度タスクの実験画面サンプル

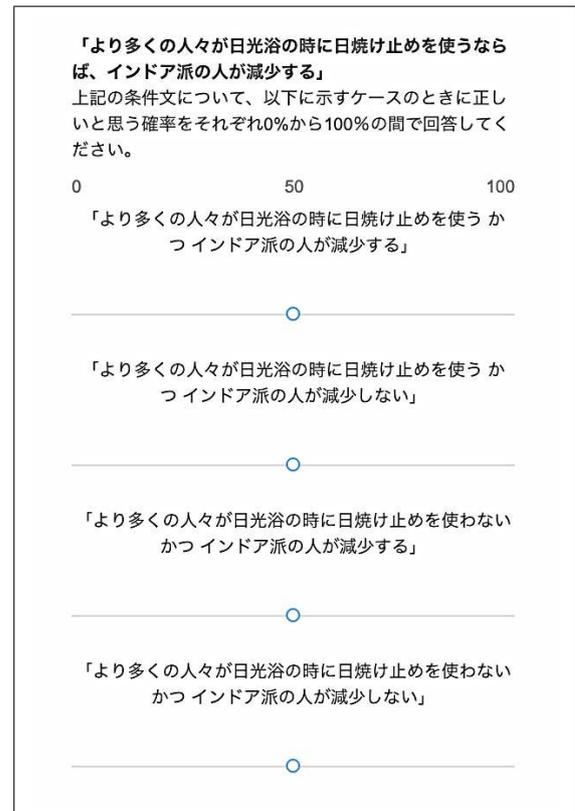


図 4 Jeffrey table タスクの実験画面サンプル