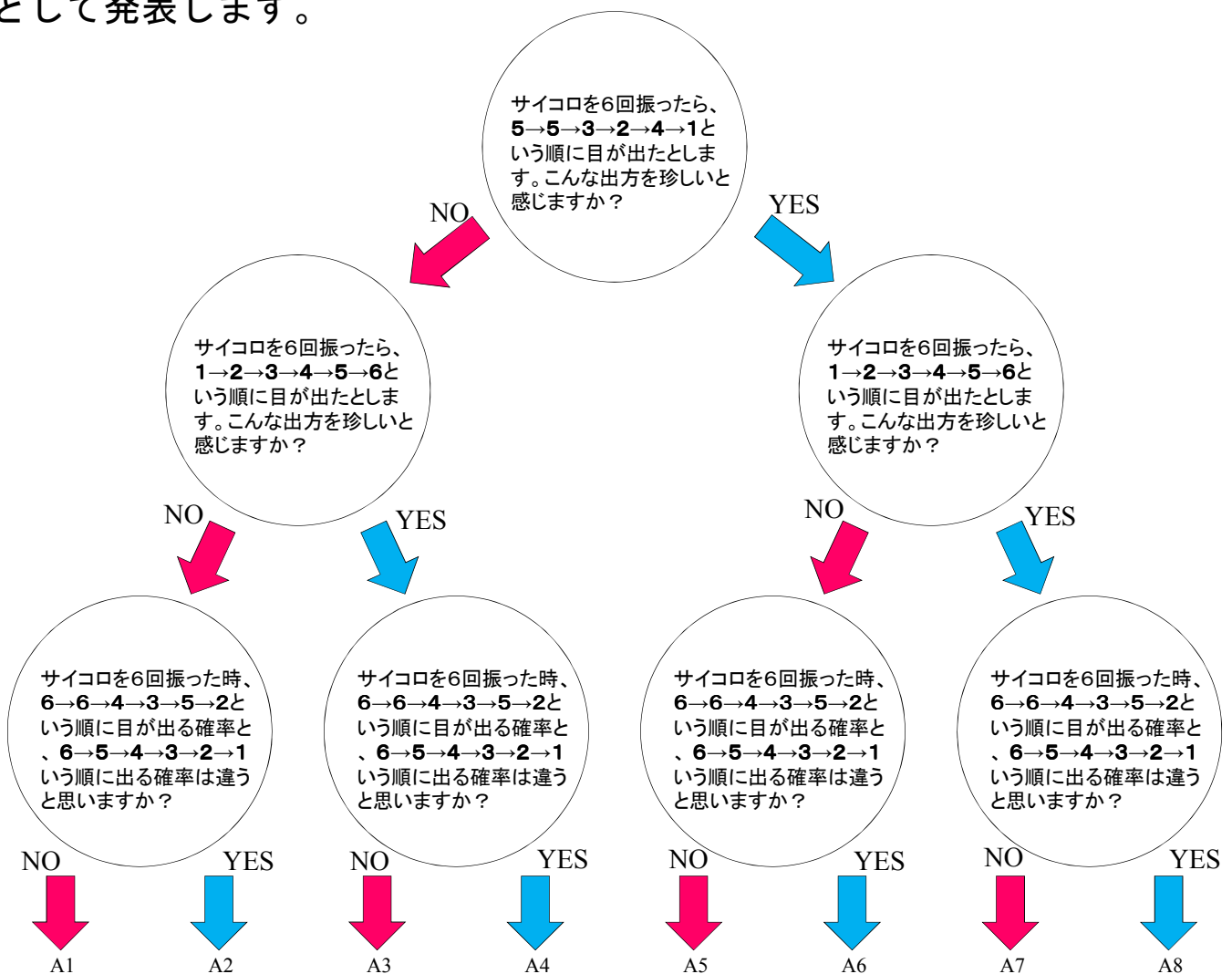


「珍しい」とは？

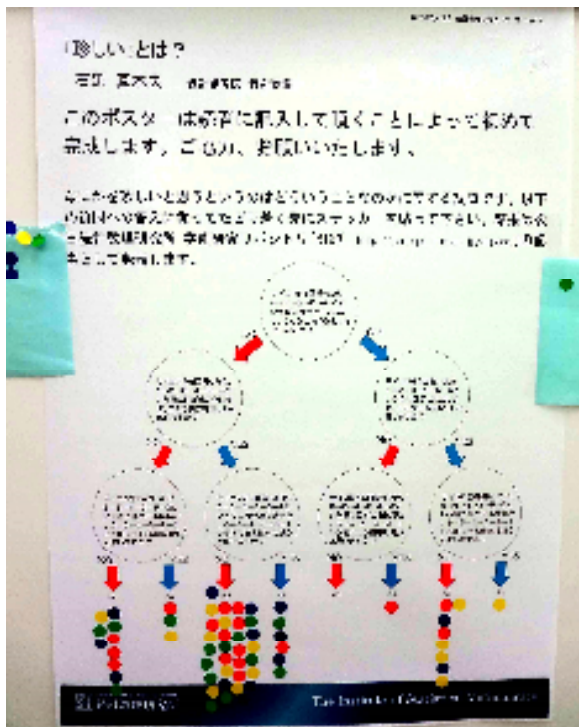
石黒 真木夫 統計思考院 特命教授

このポスターは読者に記入して頂くことによって初めて完成します。ご協力、お願いいたします。

なにかを珍しいと思うというのはどういうことなのかに関する調査です。以下の質問への答えに従ってたどり着く先にステッカーを貼って下さい。結果は後日統計数理研究所 学術研究リポジトリ[RISM <http://ismrepo.ism.ac.jp/dspace/>]の記事として発表します。



調査結果



選択肢	ポスター	メール	計
A1	10	3	13
A2	3	0	3
A3	32	26	58
A4	6	6	12
A5	0	0	0
A6	1	0	1
A7	8	5	13
A8	1	1	2
合計	61	41	102

上図左側は6月19日の17:57に撮ったポスターの写真。ポスターの前に筆者が立っていたのは11:50~12:30。既にA3にステッカーが集まる傾向が明確になっていてその意味合いについてのdiscussionが楽しかった。

上図右側の表の「ポスター」欄にポスターのステッカーを数えた結果をまとめた。「メール欄」は5月26日~6月24日の間に知人に送ったメールへの回答の集計である。オープンハウス来場者数が242との発表があった。所員による投票も含まれているので来場者の1/4にステッカーを貼っていたとは言わないが、当たらずといえども遠からずだろう。

解析

前ページの表は3つの質問

- ◆Q1:サイコロを6回振ったら、5→5→3→2→4→1という順に目が出たとします。こんな出方を珍しいと感じますか？
- ◆Q2:サイコロを6回振ったら、1→2→3→4→5→6という順に目が出たとします。こんな出方を珍しいと感じますか？
- ◆Q3:サイコロを6回振った時、6→6→4→3→5→2という順に目が出る確率と、6→5→4→3→2→1という順に出る確率は違うと思いますか？

への回答の分布を4重クロス表の形に書いたものとみなせる。

表1. 調査結果

					調査法			
					ポスター	メール		
Q1	NO	Q2	NO	Q3	NO	10	3	
				Q3	YES	3	0	
		YES	Q2	NO	Q3	NO	32	26
					Q3	YES	6	6
	YES	Q2	NO	Q3	NO	0	0	
				Q3	YES	1	0	
		YES	Q2	YES	Q3	NO	8	5
					Q3	YES	1	1

Q1変数とQ2変数を組み合わせて、Q1とQ2への回答が一致している場合に「YES」、一致しない場合に「NO」を返す新しい変数Q12を定義すると、Q12変数、「調査法」、Q3変数の関係に焦点をあてた次の表が得られる。

表2.

				Q12(一致?)	
				YES	NO
調査法	ポスター	Q3	NO	18	32
			YES	4	7
	メール	Q3	NO	8	26
			YES	1	6

「感性」と「知識」

Q1とQ2(ひいてはQ12)は数字の列を見てどう感じるかという「感性」に関する問いであり、正解はない。これに対し、Q3は確率に関する知識を問う問いであり、あとで詳論するが、NOが「正解」である。この知識とサイコロが5→5→3→2→4→1という出方をしたときと1→2→3→4→5→6という出方をしたときの感じ方がどう関係するかが気になる。これを見るために表2を「つぶして」表3を作る。

表3.

		Q12(一致?)	
		YES	NO
Q3	NO(正解)	26	58
	YES	5	13

確率論の知識の有無の感性への影響を見るために、Q3で条件づけしたQ12の条件付き確率の形に直して次の表を得る。

表4.

		Q12(一致?)	
		YES	NO
Q3	NO(正解)	31%	69%
	YES	28%	72%

Q3とQ12が見事に独立である。

この結果を「確率」と「珍しさ感覚」は無関係と読むべきなのだろうか？

discussion

質問 Q3 に対する正解は NO であるとした。質問をちょっとだけ変えて再録する。「変更」が質問の本質を変えていない事に注意されたい。

- Q1:サイコロを6回振ったら、 $5 \rightarrow 5 \rightarrow 3 \rightarrow 2 \rightarrow 4 \rightarrow 1$ という順に目が出たとします。こんな出方を珍しいと感じますか？
- Q2:サイコロを6回振ったら、 $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6$ という順に目が出たとします。こんな出方を珍しいと感じますか？
- Q3:サイコロを6回振った時、 $5 \rightarrow 5 \rightarrow 3 \rightarrow 2 \rightarrow 4 \rightarrow 1$ という順に目が出る確率と、 $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6$ という順に出る確率は違うと思いますか？

Q1とQ2では「こんな出方」という書き方をしている。Q1を読んだ人は、問われているのは「 $5 \rightarrow 5 \rightarrow 3 \rightarrow 2 \rightarrow 4 \rightarrow 1$ 」という特定の出方にかぎらず、これに代表されるある種の一群の出方についての質問と読むことができる。

これに対して、Q3では「 $5 \rightarrow 5 \rightarrow 3 \rightarrow 2 \rightarrow 4 \rightarrow 1$ に代表される一群の出方」という読みかたはできない。確率論の知識を持った人はこの質問にNOと答えざるを得ない。

しかし、Q1とQ2に答えるとき人は、無意識に

- Q3':サイコロを6回振った時、 $5 \rightarrow 5 \rightarrow 3 \rightarrow 2 \rightarrow 4 \rightarrow 1$ というような順に目が出る確率と、 $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6$ というような順に出る確率は違うと思いますか？

という質問に答えているのではないだろうか？この質問であつたら、確率論の知識を持った人も「YES」と答えられる。「 $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6$ に代表されるような何らかの規則性をもった事象」の確率と、「 $5 \rightarrow 5 \rightarrow 3 \rightarrow 2 \rightarrow 4 \rightarrow 1$ に代表されるような特に規則性を持たない事象」が異なる確率を持っていて何ら問題はない。Q3'への回答とQ3への回答が独立であっても不思議ではない。

人はものごとを「特別なオーラを放つ事象」と「その他大勢事象」に分けて認識しているのではないか。そして「特別なオーラを放つ事象」でありかつ「出会う確率が低い事象」が「珍しい事象」になる。

人は「何らかの規則性をもった事象」にオーラを感じるようである。「規則性」を持った現象の研究が法則の発見につながった成功体験の蓄積があるからだと思われる。

メールでの回答にコメントがつけられているものが多かった。中に「結果が公開されるのなら、A1, 心情的にはA4」と書かれたものがあつた(ここでの「解析」ではA1を採用した)。この例に限らず、データ全体の分布の裏に「心情」と「信条」の衝突が透けて見えるように思われる。

ポスターによる調査について

表2から調査法によって条件づけたときの条件付き確率も求められる。

		Q12(一致?)	
		YES	NO
調査法	ポスター	36%	64%
	メール	22%	78%

この結果は回答が調査法によって違うことを意味している。詳述はしないが、情報量規準によるモデル評価は「調査法」の影響が統計的に有意であることを支持する。調査対象の差かもしれないが、他人の回答分布を見ることの影響があるに違いない。

ポスターによる調査のデメリットであるが、調査のそもそもの目的が「珍しさ」なるものをデータに基いて考えてみましょうということであり、結果を見ながら討論できるメリットは大きい。