

予測市場の公共利用の方法と限界

岡野 匡志

On the public use of prediction market

Tadashi OKANO

Abstract

In this paper, we study conditions for proper use of prediction markets. The prediction markets can provide an accurate probabilistic forecast when it is appropriately operated. However, the prediction market has not been used to support public decision-making. One of the reasons is that its output has not been understood enough to be credible. Therefore, we analyze the compatibility between this method and targets. Forecasting methods are used to overcome the "predictive difficulty" with targets. However, only a few researchers associate types of methods with "prediction difficulty." Therefore, firstly we examine comparative experiments of prediction markets and other methods. In this work, we approach the difference in the nature of "prediction difficulty" with the prediction target. According to this research, "predictive difficulty" is made up of three factors. The first case is that the information that is used for inference is scattered. Second, there are complex relationships between variables, which consist of different situations. Finally, there are chaotic uncertainties. Prediction markets collect scattered information, and we can make a collective evaluation of uncertainty. However, when a situation is involved with "complexity," this performance is reduced. Discussion combines scattered information with complexity. Experts show excellent performance when there is complexity that can be structured. However, these methods cannot properly handle multiple subjective probabilities, and we suggest the appropriate role sharing of forecasting methods.

I. はじめに

I. 1. 意思決定と予測

政策形成にとって集団的予測の意味とは何か。計画立案を伴う政策形成にとって何らかの予

測は必要である。集団的な意思決定を行うには、その前提となる状況の推定は不可欠である。適切な計画は適切な予測を要する。

本稿では、集団の意思決定のための予測に関わる認識の集約方法において予測市場の意義と限界を論じる。将来の政策課題の意思決定に際して、意思決定者は、そのことがらに関わる何らかの「予測」を必要とする。予測への信頼や理解は、集団の合意形成に大きな影響をもたらす。政策課題のうち、特に以下の問題では、信頼性の高い予測は必須のものである。

- ・大規模開発プロジェクトでの、実施後の経済効果や社会への影響
- ・科学振興政策では今後の重点分野や期待されるイノベーションの見通し
- ・自然災害や原子力など科学技術のリスク、遺伝子組み換えなど未知の技術の産み出す利益や福利とリスクの受容を巡る問題
- ・外交や安全保障など、実験のしようのない問題

今までの（特に国内の）既存の政策的な意思決定に関わる予測は、災害予測や開発計画の経済効果を主な対象とし、統計的予測を元に専門家による予測と討議を加える形で行われてきた。そして、意思決定者を含む「素人」は専門家の判断を受容するか否かの決定のみに関わりえた。同時に多様で異質な人々の予測をめぐる判断の集約には様々な制約が存在する。例えば後述する「集団思考」(groupthink) 化を回避する必要がある。

こうした予測の集約の困難さ及び関与主体の限定を回避するための手法の1つとして予測市場がある。予測市場は、特定のイベント（出来事）に関する「多様な知識レベルの参加者」の予測を「市場取引による相互作用」を通じて集約する方法である。本稿では、予測市場を政策形成に利用できる条件を考えるために、他の手法との比較を通じて検討を加える。

1.2. 予測研究の動向

近年、社会における予測のあり方が問われている。第一に、多くのプロジェクトで甘い経済効果や損益予測が行われ、それを再検討する適当な仕組みがないこと、社会の不透明が増し、専門家の間でも予測についての合意が取れにくくなったこと、第二に、リスク受容やテクノロジーアセスメントなどで専門家の予測をそのまま「受容」するのではなく、様々な不透明性に社会がどう予測を受容し扱うか市民参加が求められるようになり、予測についてもその生成プロセスの透明化やプロセスへのアクセシビリティが求められるようになったこと、第三に「専門家」の予測の精度が特別に高いものではないことが知られるようになったこと、などの理由による。特に第三の理由は、民主主義社会におけるテクノクラート支配に対する批判に通底するものとなる。

現在、政策形成において一般的に利用されているのは、その問題に長年職業的に関わってきた専門家による予測である。専門家は、状況を定義する能力や、関係する知識の保有量については一般的に素人を大きく上回ると考えられている。また特定の知識ドメインに関連する情報(シグナル)¹⁾を解釈し、他の情報との関係を定義することに優れる。

特にデータが不十分か、または新規性・独自性のある単純な経験的判断が難しい状況や、将

来の不確実性に関する見積もりが必要になるとき、多くは専門家の判断が求められる。

しかし専門家の判断にも大きな誤算がある。政策立案にあたって、社会予測を行う際の困難は、ランダムな要因が全体の方向を決めてしまうことにある。認識可能だが実際に考慮しづらい要因も存在する。例えば氷山効果の把握の困難はその最たるものである。社会の大きな変動を説明するのに「氷山効果」がある。氷山の下部が温かい水で溶けることで重心が上に移り、氷山がひっくり返って新しい均衡が生まれる。社会の安定を保障しているコンセンサスの中でも見えない部分の「風化」は目立ちにくい、ある時点で均衡が逆転し転覆（例えば「革命」）するに至ることがある（宮脇 2008）。

Tetlock（2005）によれば、専門家の予測の精度は他の予測に比べて高いというわけではない。20世紀末に起きた主要な国際政治・経済上の事件（ソ連崩壊や湾岸戦争、日本のバブル経済など）を予測したデータを分析したところ、彼らの専門分野、地位、経験に関わらず予測力の優越は確認されなかったとしている。専門家の平均的な予測の成績は、わずかに偶然を上回る程度であった。また専門家が起こりえないとした事象の15%が実際には発生し、逆に必ず起きるとした事象の25%は実際には発生しなかった事を示した。特に新規な問題状況に対しては専門家の予測は素人とほとんど差がないという結果も示されている（Armstrong 1985, Camerer & Johnson 1991）。

また、経済分野において Makridakis et.al（2009）は、リーマンショックを専門家がどのように予測したかを検討し、予測すべき対象の不確実性の大きさと予測能力の限界の自己評価を適切に捉えられていないこと、にも関わらず、どちらの例でも専門家は自分の判断を過信していたことを示した。

1.3. 本稿の問題

こうした予測に対する環境の変化を受け、ここ10年ほどの間に集団的な予測生成法である「予測市場」への注目が大きくなった。むしろ現時点では予測市場はその期待の大きさに反して、政策形成への利用はほとんど進んでいない。しかし、予測市場は他の集団的予測手法に比べ、以下の大きな利点がある（岡野 2015, 2016）。

- ・不特定多数の参加が可能。必ずしも専門知識の持ち主である必要はない
- ・参加しうる人数に上限がない。むしろ数が多い方がパフォーマンスは向上
- ・オンライン化が容易。むしろオンライン以外では実施しにくい
- ・時間的な拘束がない。

そのため、予測市場の信頼性や利用可能性を高めることは、人々の予測形成への参加を支援する上で有益であると考えられる。

本稿では予測市場を「予測される対象の性質」とそれとの「相性」から再考察する。同じ予測手法でも、それがどんな状況にも適合するとは限らない。予測市場は一時過剰な期待をうけ、多くの企業で導入が為されたものの、企業予測市場のインタビュー調査（Wolfram 2015）に

依れば、調査無理解に基づいた不適切な利用もまた多かったとされている。

そこで、本稿では近年の予測研究を踏まえて「予測研究のアプローチ」と「手法の比較」を整理し、それを踏まえて予測市場の適切な利用について考察する。

Ⅱ. 政治的問題と予測市場

Ⅱ.1. 予測研究の動向

予測手法の研究は、冷戦時代に米国で軍事技術を始めとする将来の科学技術の発展の方向を見定め、科学技術政策を支援する目的で始まり、その中でデルファイ法やシナリオプランニングなどの予測手法が開発された。

日本においても科学技術庁（当時）、後に科学技術政策研究所を中心に、1971年より継続的に将来の技術発展に関する予測調査が行われている。日本の予測調査の特徴としては、専門家によるデルファイ法が広く用いられていることが挙げられる。

これに対し、1990年前後からこうした統計データと専門家の討議による予測研究の流れとは別に「予測という認知行為」や「主観的確率の認識と集約」という視点から集合的な予測生成を捉え直そうとする研究が始まった。

Ⅱ.2. 予測市場とは

1988年、アイオワ大学のグループが「予測市場（prediction market）」を開発した²⁾。同年にアメリカ大統領選挙を対象として実施し、大手メディアの世論調査を元にした得票率予測（平均誤差1.9%）よりも正確な予測（誤差0.1%）を示し注目されるようになった³⁾。

予測市場は、主観的確率の抽出と集約を「市場取引」に見られる情報（洞察）の集約機能を応用して行うものである。問題となる状況（大統領選挙など）において予測される結果（「オバマが大統領に当選する」など）を証券に見立て、実際にその予測どおりの事象が生じた場合に一定の配当がなされる「予測証券」を、専用に作られた市場を介して参加者に取引させる。参加者は予測証券の市場価格と自身の主観的確率を比較することで、主観的確率分布の自覚を促進し、また利得のためにより正しい予測をしようと新年の更新を随時行う。

この手法は多くの実験で高い予測的中率を示し、2001年には米国防総省の国防高等研究計画庁（DARPA）が政策分析市場を実施した⁴⁾（Hanson & Hertwig 2005）。このプロジェクトには、テロ発生の可能性を評価するため、テロが実際に発生した場合配当がなされる証券が含まれていた。それゆえ議会で問題視されるにいたって閉鎖された。現時点で、このプロジェクトは、国家の政策立案に直接利用することを目的とし、国家機関のプロジェクトとして公開され運営された唯一の予測市場である。こうした試みが示すところの、人々の洞察の集約の効果はJ. Surowiecki（2004）によって“Wisdom of Crowds”として纏められた（岡野 2015a）。すなわち、専門家を含めた議論によって正確で高度な集合的推測が可能になるのではないこと、また素人集団と専門家集団の集合的推測の正確さには差はないこと、それどころか素人集

団の集会的推測の方が正確である場合も多々あるという主張は、社会に驚きをもたらした。

II. 3. 予測市場の利用

実社会での意思決定に予測市場が利用された例は既に数多い。予測市場が一時期過剰とも思える注目を受けた背景には、それが「専門家以上の予測精度を示しうる」「信頼しうる代替予測を非エキスパートの集団で作る」と期待させたことが大きい。企業による社内予測市場でも、新製品売上や将来市況の予測などで社内の専門チームの成績を上回るなど、「専門家以上のパフォーマンス」を示した例がいくつも報告されている。適切な環境で適切な対象に対して実施された予測市場の成績は安定して良好だとされる。予測市場で予測されたイベントの発生確率と実際の発生率を比較するとほぼ一致している⁵⁾。

集団や組織の意思決定への応用も 2013 年のマッキンゼー調査 (McKinsey & Company, 2013) で、対象となった企業の 7% で利用経験があるとされ、かなりの広まりを見せている。しかし、公共的な政治的意思決定に直接利用された事例は上記を除いて公表されているものはまだ無い。

ところが予測市場の予測は、大きく外れることも場合によっては無いわけではない。あくまで予測市場は参加者から予測を誘発し、その予測を相対的評価（重み付け）して集約する機構であって、本来存在しない知識の代替となるわけではない。また失敗が不適切な対象や運用の結果だったとしても、原因を確定する方法は未だ不十分である。

そして予測市場の出力を「どれほど信用するのか」を判断するための方法もまだ無い。このように予測市場は集団が意思決定のために用いるツールとしては、まだいくつかの課題を残しているといえる。

II. 4. 予測トーナメントと予測手法の検討

米 ODNI（国家情報戦略オフィス）が管轄する知能高等研究計画活動（IARPA）においても 2010 年 6 月から 2015 年 6 月まで、「リスク評価」の手法の研究のプロジェクトとして Aggregative Contingent Estimation (ACE) Program⁶⁾ が始まり、その中で（統計モデルではなく）認知または集約条件の異なるグループ間での予測力を比較する予測トーナメントが 2011 年から実施された。

ACE の目標は、広く分散した多くの分析者の判断を集約することで、国際問題に関する事件（イベント）に生起についての正確でタイムリーな確率予測を生成することだった。このトーナメントでは、いわゆる「群衆の叡智」の利用による「クラウドソーシングインテリジェンス」が特定専門家のグループの予測能力を上回るのかを検討することが目的とされた。

トーナメントには 5 つの大学・研究所チームが参加し、各チーム合わせて数千人の参加者（専門家、大学生、一般含む）を募集した。彼らは以後 4 年の間にアメリカの安全保障に関する地政学的もしくは経済的な約 500 件のテーマ（「欧州連合（EU）から脱退するメンバーが現れる可能性は？」「東シナ海で 10 人以上の死者が出る衝突の可能性は？」「一年後の原油価格は？」

など)について判断させた。

各チームは独自のミニトーナメントを行い、よりよい予測の誘発、収集、収集した予測の組み合わせの方法の開発を目指した。そして独自の回答の方法を用いて参加者の意見を収集し、単純平均や予測市場、それぞれ開発した「クラウドソーシング」を利用した(統計的手法、アルゴリズムとの組み合わせなど)多様な手法を用いて集約し、その結果(確率推定値)の精度を競った。

各チームの成績(予測の評価)にはブライヤースコアが用いられ、チームの集合的予測生成のために4年間で集められた全質問への全回答の総数は100万件に上った。(Tetlock, et al 2014, Mellers et al 2015, Wintle, B., et al. 2012)

II. 5. 予測市場の問題

予測市場は、このトーナメントでも用いられ、多くの予測手法と同一の問題を予測し、その予測力を比較されている。予測市場の成績はチームによっても異なるが、多くで良好ではあるが最も高い予測力を示した手法では無かった。この件について後の比較で言及するが、予測市場は現在では「正確さを証明された唯一の集団的確率予測法」ではない。

しかし完全な予測法は無い。いくつもの確率予測法が登場したからこそ、予測市場というメカニズムがどのような問題に適しているのか、予測市場の予測結果の信頼性をどのように事前に見積もるかといった問題を改めて考え直す好機であろう。予測市場の限界の自覚ができれば、逆に補完による対処(組み合わせ)や歪みの抑制の対応をすることができる。これは、予測市場の社会的な利用を進める一助になるはずである。

III. 予測市場と他の手法の比較

予測市場が社会的に有効な手法として普及するには、ただ正確な予測ができる。というだけでは意味がない。

先に挙げたように、予測市場はあるときは高い予測力を示す一方、予測生成に関係する変数が多く、予測力の期待値を見積もりにくいという問題がある。時に非常に正確な予測を示すシステムであったとしても、期待される誤差が不明であればそのシステムを社会的な責任の生じる問題に利用することは難しい。

大統領選挙やスポーツの試合以外の対象を扱ったとき、予測市場の「予測精度の安定性」はどの程度期待できるのか、既存の手法や代替手法と比べ、どんな状況でいかに優れたパフォーマンスを期待できるのか、もし予測精度が低下するのであれば、それはどのようなときに起こりやすいのか、といった問題を整理する必要がある。しかし、これらは十分に解明されていない。

まず、予測力に影響する変数群としては、①予測対象の性質、②参加者の数や能力・態度、③集約メカニズムとその中で生じる相互影響がある。このうち②③については既にある程度検

討が行われているが、①どのような状況で予測市場が他の手法より有効なのかを、個々の事例ではなく整理した知見はまだない。確かに、これを実証するには十分に統制された数多くの大規模な実験と調査が必要であり、現実的ではない。

Ⅲ. 1. 予測市場と他の手法の比較

実際に予測市場は、他の手法に劣っているのであろうか。それを確認するために、以下では同一の予測問題を予測市場とその他の手法で予測し結果を比較した知見の整理を行う。今回は比較研究のうち、特に問題状況のあり方と予測手法の関わりに触れているもの、次章で触れる「予測困難さ」と予測市場の機能を考える上で参考になるものを取り上げた。

問題の条件としては（デルファイ法などは定性調査に主に使われるが）予測市場との比較を目的とする関係上、イベントなどの発生確率（結果は実際に発生した・しないのバイナリ）、もしくは値（定量）などの検証可能なものに限っている。予測市場にもいくつかの方式があるが、ここでは最も一般的な「ダブルオークション」方式（以下 DA）⁷⁾と比較する。特殊なマーケットメーカーを利用した市場は扱わない。また、パリティミュチュエル方式（以下 PM）⁸⁾の予測市場は DA の比較対象としてのみ触れる。

比較対象とする集団的予測手法とは、討議やデルファイ法など、予測者がグループを作り、予測や状況判断（両者を併せて「洞察」とする）を何らかの方法で交換しながら集約する。それを通じて、合意かもしくは均衡による「集約的予測」を生成する手法を指すものとする。なお、予測市場の他デルファイ法やノミナルグループ法など、洞察の抽出・表明・交換が一定のルールとフォーマットの元で行われ、予測者同士の相互作用への統御が行われている手法を以下では「構造化された予測法」と呼んでおく。

Ⅲ. 2. 大規模集団による予測の比較

ここでは先に述べた予測トーナメントにおける検討を中心に、数百人以上の参加者による予測の比較を行う。予測市場は他の参加者同士のコンセンサスを必要とする手法と異なり、「均衡」を予測値として用いるために多くの参加者の利用に適している。この条件で他の予測手法より良好な予測を生成することは予測市場の価値を考える上で重要である。

Ⅲ. 2. 1. Good Judgement Project (GJP)⁹⁾

Good Judgement Project は P.Tetlock をはじめペンシルバニア大とカリフォルニア大パークレー校を中心としたチームである。このチームでは、大学院生から政治学教員や実務家に至るまで 2,000 人の回答者を募集し、総計 10 万本の確率予測を収集した。それを元に彼らは以下の検討を行った（Unger et al 2012, Atanasov et al 2014）。

- ・ 予測市場と予測投票（Prediction Poll、以下 PP）の比較。PP は主観的確率を直接抽出して集約する方法。指定された事象の発生率を回答し、それを直接統計アルゴリズムを用いて集約する。この時どのような処理をすれば PP は予測市場のパフォーマンスを上回れるか。

- ・ 回答者を数人ずつチームにするチーミングの効果。
- ・ トレーニングの効果。参加者に自分の主観的確率評価を認識するトレーニングを行った場合、予測は正確化するか。

それによれば、

- ・ 予測市場は個人・チーム世論調査の単純な非加重平均を大きく上回った。また PP で投票された予測の時間経過による重みの割引、予測者の過去の成績による重みの加減を行うと個人・チームどちらも予測精度に大幅な改善が見られたとされる。すべての修正を行うと、PP の成績はどちらも予測市場を上回った。PP では、予測されたイベントの発生確率と実際の発生率にはズレがあったが、重みの修正後は一致した。なおこの重みの修正による改善効果は予測市場では見られなかったが、予測市場ではすでに予測確率と発生率をよく一致しており、修正は却って逸脱を大きくする結果になっている。

PP は専門家同士の意見の速やかな集約、コンセンサスが不要という点で予測市場に匹敵する。

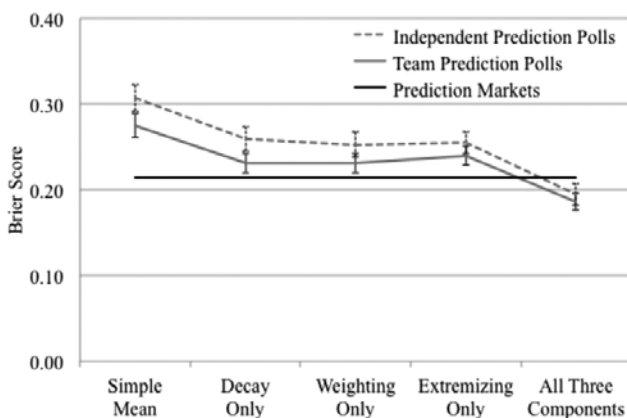


図 1：予測市場と PP の精度

[出所] Atanasov et al (2014) より引用。PP の重み付けアルゴリズム別ブライヤースコアと予測市場の比較。黒実線が予測市場の成績。破線は個人参加者の予測投票成績、灰色実線がチーム参加の成績。Y 軸はブライヤースコア (0 が最良)、X 軸は適用する処理。

加えてトレーニングの効果も検討され、個人でこうした複数の主観的確率を生成、結合、調整し、集団的予測を上回る予測を行える「Super-forecaster」の存在も主張されているが今回は触れない。

Ⅲ. 2. 2. Olson et al (2014)

予測トーナメントにおいて直接自信の確率判断を入力するインタフェースを有する予測市場を用い、三点誘発法を用いてより主観的確率の抽出効率を高め予測の正確化につなげようという試み。三点誘発法は主観的確率の上限・下限・最高値を回答させ確率判断を効率的に引き出す手法である (Speirs & Bridge et al 2010)。しかし予測結果に有意な変化は見られなかった。

このことから、予測市場は既にその機序で利用できる範囲内で十分な確率分布の抽出を行っていると考えられる。

これら予測トーナメントで行われた諸手法の比較研究は他にもいくつかある¹⁰⁾が、これらの検討は稿を改めたい。

Ⅲ. 3. 小規模集団による予測の比較

続いて、参加者が10人以下の条件での予測を紹介し比較する。

Ⅲ. 3. 1. Greafe & Armstrong (2011)

Greafe らは予測手法の比較研究を系統的に行っている研究者である。この実験では大学生227人を5人程度ずつ44のグループに分け、それらを対面式討議（FTF）とノミナルグループ法¹¹⁾、デルファイ法、予測市場の4つの手法に振り分けて、事前に調べた回答者の見積りの誤差平均（ほぼ差は無し）からどれだけ改善されたかを比較した。問題は「オーストラリアの人口」「65才以上の人口比」「健康への支出」などの知識を利用した見積り課題であった。結果は押し並べて改善（3.6～3.05）が見られ、手法の間に統計的に有意な差は見られなかったが、個々の問題ではデルファイ法が最も成績が良く、次にNGTで、予測市場は最下位のFTFとほぼ変わらなかったとされる。

その原因の一つとして、回答者間の直接の情報交換の機会が無いことが信念更新の機会の不在や推測の自己評価の不十分さにつながり、取引を不活発にした可能性が指摘されている。また、（本稿では述べられていないが）問題が確率判断の交換で回答可能なものでなく、知識の交換が必要な見積り課題であったことの影響も考えられるのではないか。これについてはまとめで再度触れる。

なお、各手法の問題ごとの改善率の分散（問題による効果の揺れ）はNGTが最も高く、デルファイ法が最も低い（安定していた）。予測市場は中間になる。参加者の満足度はNGTが最も高く予測市場が最低。また、参加者の回答の差異は予測市場が飛び抜けて大きい（信念の収束が起きていない）。

Ⅲ. 3. 2. Healy et al (2010)

情報集約において・参加人数・環境の複雑さ・異常な個人の行動は予測手法のパフォーマンスにどのように影響するのかを検討している。比較対象は（DAの他）、PM（パリティミューエル予測市場）、IP（反復投票 McKelvey & Page 1990）、MSR（マーケット・スコアリング・ルール Hanson 2003）。IPは一種のデルファイ法である。大規模な状況でのDAの有効性はすでに示されているが、ここでは最少人数を用いて異なった状況での手法の効果の比較を行っている。

回答者はメールで募集した大学生を1チーム3人ずつ各手法に分配した。問題はコインの分布の予測（人工的に統制された状況）が用いられた。これによれば、状態数の少ない単純な状

況では平均的に MSR が誤差最大。続いて PM、DA が最小であった（有意水準 10% の場合）。また、予測される項目の数に対する回答者が十分に多いのであれば DA のパフォーマンスは他より高い。しかし、項目数が多いときやイベント間に相関が存在する場合（複雑な環境下）では DA は最悪のパフォーマンスを示し、PM がそれに次ぐ。上記のような複雑な環境や参加者が少数であれば IP が有効性を示した。

逸脱としては、PM と MSR は異常な個人の予測に影響されやすい。また予測メカニズムの出力が間違った推論を呼び起こす「蟹気楼」による予測の悪化はすべての手法で見られたが、単純な状況では統計的に有意な差は無い。しかし複雑な状況では DA で最も多く発生した（IP で最小）。

Ⅲ. 4. 比較の総括

これらの他にも実験的な比較検討の知見があるが¹²⁾紙幅の関係もあり同じく稿を改めたい。

本稿では、問題状況との関わりについて比較した知見をいくつか取り上げた。予測市場は、主観的確率の抽出と分散情報の集約に優れるが、先に見た Healy (2010) でも DA の予測市場を含む各手法の成績は、単純な問題状況と複雑な問題状況で（少人数の元では）パフォーマンスが大きく変わることが示された。しかし状況の複雑性の計測はそれ自体が困難であり、予測市場に生じる困難を事前に評価することは難しい。

次章ではこれらの検討を元に、問題を取りまく状況と予測手法の関わりについて整理を進めたい。

Ⅳ. 予測の困難さをもたらす要因の整理

これらの様々な予測手法は、「予測の困難さ」を克服するために考案されたものである。予測研究では対象となる問題状況に最適な「予測法」を同定することが重要になる。予測手法の研究では、各種法がどのような「困難さ」の克服に有効で、何を克服できていないのかを個別に検討されることは多かった。しかし、それらを総括し、予測手法全体を予測困難さとの関係で位置づけた知見はない。

もちろん、予測手法が扱わねばならない「変数」は非常に多く、さらにどのような「変数」が関与してくるのかを前もって十分に確認しておくことも難しい。また、予測対象の状況にどのような「困難さ」が潜在しているのかを客観的に定義できる指標も不在である。このような理由のため、問題状況と予測システムの「相性」は、まだ曖昧なままで十分な分析も整理も行われていない。

予測市場に対する過剰な評価と批判が同時に存在するのは、こうした状況が理由の一つであろう。予測市場の適切な利用のためには、それが「どのような問題に効果的なのか」「対象となる状況はどのような『予測困難さ』をもつのか」を理解した上で、その限界を補完できる他の手法と組み合わせることが必要であろう。

そこで以下では、先の予測手法の比較を踏まえ、予測の困難さが何故生じるかを整理してみたい。

IV. 1. 予測困難さの分類

図2は、「予測の困難さ」とその克服機能を持つ予測手法の関係を纏めたものである。先の予測手法の比較の中で、以下の3つの「予測の困難さ」の原理が抽出された。その一つが、情報の分散性であり、二つめが状況の複雑性、三点目が不確実性である。

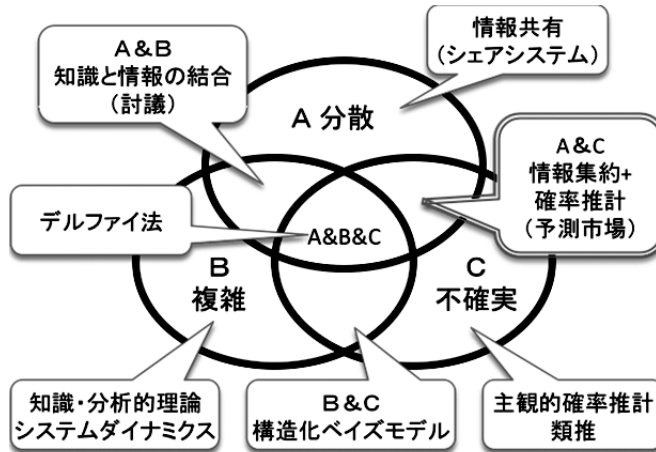


図2：予測の困難さの原理と予測手法の位置づけ

【出所】 筆者作成

IV. 1. 1. A：分散性

状況判断のための情報自体は十全に存在するとしても、一人一人がアクセス可能な情報が部分的であるなら、それぞれは不十分な情報に基づいて予測を行う事になる。選挙結果などでは有権者が誰に投票するかという主観的判断が問題となるが、これは本来個々人の内面に所属し、シグナルも個別に発せられる（全てのシグナルを一覧する視点の不在）。

この場合は、それぞれが得た情報をシェア（共有）することで全体の情報を確認して克服できる（Yang & Heck 2015）。ただし、シェアの際に発生するバイアスが十分にコントロールされていることが重要である。例えばシェアされた情報が存在しうる情報全体から見てどれくらいの「偏り」を持っているのかは（母集団自体の構成の情報も分散性を持つなら）そのままでは評価できない。

IV. 1. 2. B: 複雑性

問題となる状況を構成している各変数間の関係が複雑で、把握や予測のためのメンタルモデル構築が困難な場合。ある変数の変化がその状況にどのような影響を与えるかは、複雑性が高まると、多くの他の変数の相互関係の状態が影響してくるために認知的な負荷が高まり把握し

づらくなって直感的な予測が行えなくなる。また観測だけでは把握が不可能になる。

これには変数間関係の知識を元にして、複雑さを認識可能にするための（構造モデル化やシミュレーションなどの）手法を利用することで克服が試みられる。

いわゆる「専門家」は、この「複雑さの認識とモデル構築」に優れる。このため、分散性および不確実性の乏しい問題（必要な情報が十分に存在してすでに共有されており、想定外の事象の影響に制度的な制約がある場合）では専門家の判断が優越する可能性が高い。後述する Burgman et al (2011) の「討議による専門家のエラー低減」これと矛盾しない。

IV. 1. 3. C: 不確実性

対象となる状況そのものにカオス性や創発性が存在する、それ自体が未観測の事象で過去のデータに頼れない、偶然のアクシデントや想定外の事件の発生、観測不可能な外部からの影響などがある場合。GJP の言う縮減不能な不確実性に該当する。上記二つが（原理的には）完全な克服がなされれば確実な予測が可能になるはずであるのに対し、不確実性はどうしても確率的にしか定義しえない。

これについては（個人的不確実性を含めて）主観的確率推計での対応がなされる。また、同一の事例について不十分な経験しか存在しない場合や、分散性と複雑性による情報や知識の欠落や認知能力の限界（個人の無知による疑似的な不確実性）の際には「似ている（と予測された）」事例を利用した類推¹³⁾が用いられる。例えば選挙予測であれば、ある事件が結果にどのように影響するかを、周囲の反応や過去の「類似的」事例、自分自身の反応からの推量で行う。

予測を行う場合、観測した変数が「経験上のどの変数に相当するのか」を同定し、適切な状況判断の枠組みを選びださなければならない (Spence & Brucks 1997)。しかし、本当にその同定が妥当なのかは一概には言えない。変数の同一性が表面的には高くとも、その背景にある構造が近似しているとは限らないからである。このようにあるフレームで行った判断がどれくらいの「妥当性」を持つのか、逆にどれくらい割り引くか、の判断もまた不確実性を持つことになる。

この分類は、あくまで問題の明確化のための理念型である。純粋な「分散性」「複雑性」「不確実性」のどれかだけを持った問題状況は実際には存在せず、実際には3タイプが濃淡の差はあれ混交していると考えられる。

IV. 2. 集団的予測手法と困難さの原理

いわゆる「専門家」は、この「複雑さの認識とモデル構築」に優れると思われる。このため、分散性および不確実性の乏しい問題（必要な情報が十分に存在してすでに共有されており、想定外の事象の影響に制度的な制約がある場合）では専門家の判断が優越する可能性が高い（これについては次章で後述）。

対して、予測市場は、元々以下の特徴を持つ。

- ・市場の「分散した情報の集約機能」に期待した機構である。
- ・個人々の「主観的確率分布」の相違によって機能する。

これはそれぞれ「A: 分散性」と「C: 不確実性」に該当する。

先の事例のうち予測市場が安定して専門家以上の予測精度を示している典型的な事例は「選挙結果」や「新製品売上」など、分散性もしくは不確実性が強いものである。「選挙」は情報（有権者の意思）は分散しているが、制度的にも複雑性はそれほど高くはない。「新製品売上」も同じく消費者の意思および部署の異なる参加者の持つ情報は分散、かつ「新しい提案に対する購買意欲の発生可能性」は不確実性（観測し得ない外部）に当たる。

逆に、予測市場は複雑性の高い状況に対してはしばしば予測力が低下する。これは複雑性を低減させたり、知識の妥当性を検証するための機序が（討議などと異なり）予測市場の中に存在しないことが関係していると考えられる。これはあくまで「機序の不在」であり、参加者の資質や努力によって対処ができない訳ではない。しかし、適切な対処が可能であるかどうかは系統的に保証されないことは確かであり、これが予測市場の予測力が時に逸脱を起こす原因の一つと思われる。

IV. 3. 他の複合領域に適した予測手法

A 分散と C 不確実の重複領域においては、予測市場が「分散」した情報に基づいた主観的確率推定を集約することで A と C を接続して予測生成をしている。ならば、他の3つの重複領域（分散 + 複雑、複雑 + 不確実、及び三要素の重複）ではどのような手法が対応するのだろうか。ここでは、手法と領域の対応は、手法の主要な機序に「その困難さを解消する働きがあるか」によって分類した。

IV. 3. 1. 分散 + 複雑（A & B）

「分散 + 複雑」状況は、状況判断に関わる変数の値が分からない欠落が多く、情報と変数の結びつけも（知識の分散のために）難しい状況である。個人々は分散したシグナルを入手しても、複雑性のためにそれがどのような変数のデータかを同定できず、その影響を解釈できないためにシグナルが「ノイズ」として扱われてしまいやすい。このため、分散した情報を複雑な構造の中に結びつける作業が必要となる。これがいわゆる「討議」にあたる。

「討議」では、他の参加者とそれぞれが持つ情報およびその情報をどのように解釈するか（状況と結びつけるか）についての意見を交換する。この交換を通じて状況の複雑性に対処している。討議を通じて、グループに潜在している情報（要素の状態についてのシグナルやデータ）と知識（情報と要素の関係）を抽出し結びつける。

このように、討議はこの状況に適した手法であるが、話し合いを通じて不確実性を評価する事は難しい。不確実性の見積もり問題では討議の成績は高くない。この原因としては、各自の主観的確率を適切に集約する機序が不在であること、対面討議では「リスク・コーシャスシフ

ト」などの歪みが生じやすいことなどが挙げられよう。

既に何らかの確率判断が下された不確実性であれば「知識」の一つとして扱えるが、不確実性の評価は「熟議」の再評価を行う上でも重要であろう。

IV. 3. 2. 複雑+不確実 (B & C)

複雑性と不確実性を同時に持つ問題（例：原発事故の予測など）では、状況に含まれる変数間の関係だけでなく、その不確実性が、他の変数の不確実性にどのように関わっているのかを把握することに困難さがある。

これへの対処としては「C不確実」な状況で用いられた方法を構造化させた手法を用いることが適当であろう。ここでは2つの手法に言及するが、一つが構造化されたベイジアンモデルによる手法である。まずそれぞれの主観的確率を抽出して、条件付き確率のネットワークを構築し、それを用いて予測を行う。

もう一つが、構造の類似（類推）によって不確実性と複雑性（の一部）を代替する構造化された類推法（半構造化アナロジー）（Green & Armstrong 2007）である。参加者は、最初にできるだけ多くのアナロジーを提案するように求められる。その後、予測とそれぞれのアナロジーと問題との類似性の評価を生成する。本稿は予測市場の再評価を目的としているため詳しくは触れないが、Nikolopoulos et al (2015) によれば、構造化された予測法を用いないグループに対し8.4%の精度向上が見られたとしている。

IV. 3. 3. 分散+複雑+不確実 (A & B & C)

デルファイ法は、専門家の意見調査にも用いられるなど、「意見への賛否の理由の交換」共有機能もある程度有しているため、（ここでは対応する機能の有無を基準とするため）ここに配置した。とはいえ、共有の機能としてはそれほど高くない [デルファイ法の分散対応力→行方不明]。均衡によって集約する予測市場と異なり、意見交換と理解を通じたコンセンサスによる集約である。

「デルファイ法」では、永らく科学技術の将来発展などの専門性と不確実性の強い問題に適用され効果を示してきた。先の Nikolopoulos et al (2015) でも、デルファイ法は54%の精度向上を示している。

デルファイ法による予測生成ではお互いの因果や判断の根拠の交換と予測シナリオの相互評価をするが、情報の分散の解消自体は目的にしていない。お互いが既に（相互検証するための）十分な知識とリサーチに基づく情報アクセスを持っていることを前提に、それでも残る不確実性とそれが状況全体にどのように影響を及ぼすのかを検討するものである。

IV. 4. リスクコミュニケーションと討議（熟議）

リスクコミュニケーションとは、社会を取り巻く様々なリスクに適切に対処するために、リスク発生前に問題状況を定義し、その理解を進め、対処のための方策と協力関係を構築するこ

とである。リスクコミュニケーションは意思決定や合意形成に向けて行われるが、あくまで、そこに至るデータ、知識、意見、判断、価値観の集約と検討を目的とした活動である（これは集合知の対象領域同じ）。また熟議論もリスクコミュニケーションを主要な対象としており、現代的な公共性の中心的問題の一つとなっている。

リスクコミュニケーションの対象には「リスク状況の認識」と「それを受容し、統御するための集団的規範形成」の2側面がある。このとき後者のあり方は前者に依存する（逆である場合もありうるが、リスクコミュニケーションにおいては、それは一般的に望ましくない）。

この「リスク状況の認識」の困難さは、すべての予測困難さが大きく存在する状況「A+B+C」の領域であることにも関わるのではないと思われる。討議を通じたリスクコミュニケーションの中でリスクの確率的評価を行う事は、ここで整理したように実際には難しい。

討議は図ではA+B領域に適した手法である。（予測市場を含む）C領域に対応した機序を持つ手法を適切に併用することができれば補完しうるかもしれない。これは今後の課題としたい。

V. 予測と過信

次に、予測者の過信の問題を整理する。予測研究において、過信は重要なテーマである。過信があると、自分の判断に含まれる歪みや逸脱を認識し適切に修正できなくなるにも繋がる。予測市場は「自分の判断への投資」のステップを導入することで、自分の意見への信頼性への客観的な評価を誘導し、適切な「均衡」の生成を支援している。過信・もしくは自信喪失はそれが一部であれば本人の損失となるだけであるが、予測値を直接歪める諸々のバイアスとその性質に関しては稿を改め、ここでは2つの過信を検討することにした。1つが他の人々が自分と同見解であることで生じる「社会的効果による過信」であり、もう一つが、自分の知識への信頼に起因する「専門家の過信」である。

V. 1. 社会的効果による過信

人間は複数の他者が共有している意見を正しいと感じ、同調を起してその意見への疑念が消えるという強い傾向を持っている。これは自分と集団の意見が近かった場合には特に顕著で、この性質はリスクシフトなどの集団思考を引き起こす原因となっている。これは集団で合意形成を図る際に関係してくる「歪みの原因」である。予測においても、他者の予測が自分のものと近かった場合、または多くの他者の予測値を自分の予測の参考にできた場合には確信度が高まることが確認されている（Lorenz et al 2011）。

- ・ 他者の判断の傾向を知りうる場合、同調による歪みと歪んだ予測（見積もり）への過信が生じる。
- ・ 個々の回答には大きなエラーが含まれていてもそれを平均化するなどである程度制御しうる。しかし、こうした社会的効果による歪みは集団内の相互作用で強化し合い、エスカレー

トしやすいため、その集団内では平均化などの方法では制御しきれない。

V. 2. 専門家の過信と討議

Tetlock (2005) らが示した専門家の過信の原因についてはまだ解明されていないが、いくつかの仮説が提示されている。

- ・ 1つが自分の依拠する知識ドメインの過大評価である。
- ・ 2つ目が自己の状況判断の枠組み（フレーム）に関連する情報とそうでない情報への感度の差である。
- ・ 3つ目が複数の情報源の重み付けの失敗である。

Burgman et al (2011) は、医学、疫学、動物と植物の健康、生態系や保全生物学などの専門家から、関連する教育をある程度受けただけの多様な知識レベルの人々を集め、過信の制御と予測の精度向上に関する実験を行っている。

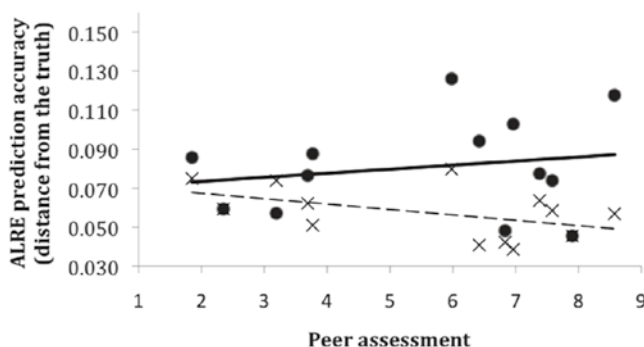


図3：専門性と討議による予測見直し

[出所] Burgman et al (11) より引用。X 軸は専門性のピア評価（大きいほど専門性が高い）、Y 軸は予測誤差。●は第一回目の予測誤差、×は討議後再予測した際の予測誤差

実験参加者は最初に関心とお互いの専門知識のレベルを相互に評価した（ピア評価）。この評価は公開されない。次に専門分野に関連する推定問題を出し回答を求めた（第一回の回答）。最後にグループ討議を行って同じ問題に再び回答させた（第二回の回答）。最初の推測値の精度（実線）は、ピア評価が高い参加者の方が低い参加者よりも推測の精度は僅かであるが低い。また、ピア評価が高い参加者間のエラーの散布も低い参加者よりもわずかであるが却って大きい。

しかし、討議を経た2回目の推測（点線）では、ピア評価の高い専門家の精度向上（エラー低減率が高い）のに対して、評価の低い、専門性の低い参加者の精度は大きくは変化していない。ここで見られた変化は、他者の視点から問題を捉え直し、クロス評価してその視点の価値や効果を評価することで自分の状況認識モデルの再構築が行われたためと考えられるが、評価の高い専門家は適切な（相互に専門的判断の妥当性を適切に評価しあえる）環境下ではその修正を（平均すれば）適切に行うことが示唆されていると考えられる。

V.3. 「複雑な状況」の構造化可能性

問題状況がどれだけ「構造化した（秩序だってモデル化しやすい）捉え方を許容するか」という問題は専門家と素人のパフォーマンスに大きく関係するとされている（Spence & Brucks 1997）。専門家がモデル化できていない新しい状況を予測する場合にはパフォーマンスが低下する（Chi 2006, Kardes 2006）。また秩序性が非常に低い問題状況や、シグナルが極めて乏しい問題であれば、「ランダムよりは少し秩序だった」程度の捉え方が限界である。

図4は状況をどれだけ構造化して捉えうるか（X軸）と、専門家（Experts）及び素人（Novices）の予測判断のパフォーマンス（Y軸）を表している。図の左側が十分な構造化可能（well-structured）、中央が不十分なもしくは不適切な形で構造化（ill-structured but structurable）。散乱した複雑性のために体系性を誤認しやすい。右側が構造化の不在（un-structured）になる。

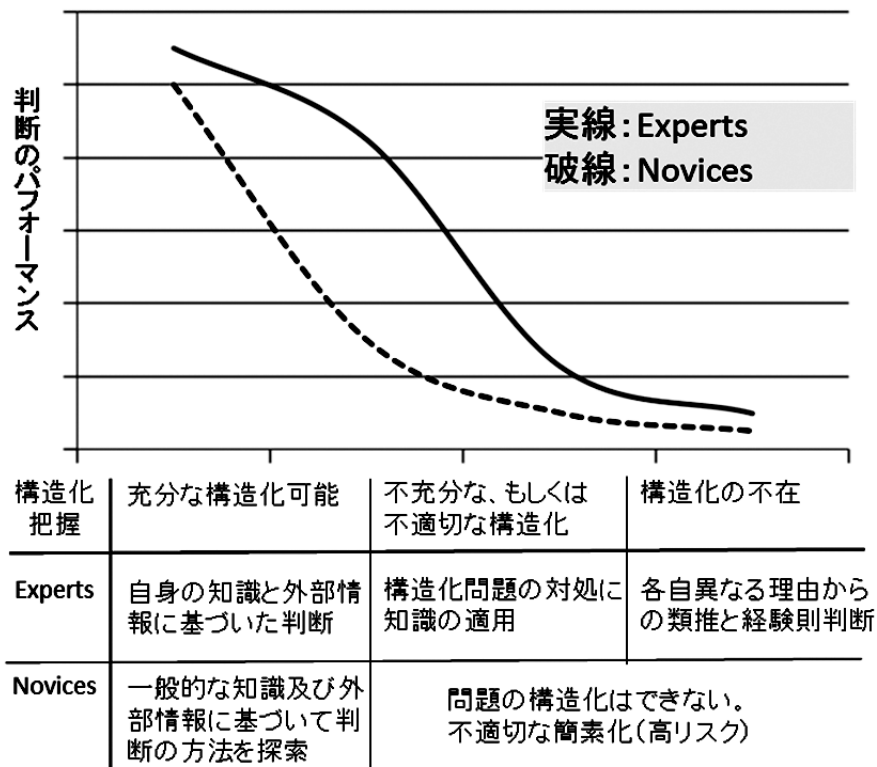


図4：状況の構造化とパフォーマンス

【出所】 Spence & Brucks (1997) を元に一部修正し再作成。

このうち well-structured であれば、十分なシグナルを受信して、適切な知識（予測モデル・ルール）に代入することで素人でも精度の高い予測が可能である。逆に un-structured であれば、専門家も素人と同様に類推とヒューリスティックに頼った予測しかできない。両者のパフォーマンスに大きな差が生じるのは ill-structured but structurable である。この範囲では、他の知

識や経験からの補完と仮説生成が必要であるが、素人はそれが困難である。ただし、自分の予測パフォーマンス以上の過信を専門家が抱いてしまいやすいのも未検証ではあるが、この領域の可能性が高いと思われる。

ill-structured but structurabile の問題状況において、予測市場のパフォーマンスと参加者の知識レベルの検討は予測トーナメントでも行われてはいるが、まだ充分には分析はされていない。これは先の専門家の討議の効果が成り立たない「相互評価可能な特定領域の専門的判断」だけでは対処できない課題、ピア評価の低い参加者を大量に含む状況での比較も合わせ、予測市場研究の今後の課題であろう。

VI. おわりに

本稿では、状況に潜在する「予測の困難さ」と予測手法の関わりについて検討した。対象となる問題状況にどのような予測手法を用いるかの判断には、

- ・ 問題状況の予測困難さの性質
- ・ 複雑さの「構造化」の程度：よいアドバイスの分別可能性
- ・ 参加人数：少ない場合は予測市場は困難
- ・ 参加者の専門性と多様性：討議が有効か
- ・ 合意による予測値（参加者が受容しやすい）と均衡どちらを求めるか

以上の5つの次元が存在しうるのではないかと考える。ただし条件と手法の対応はまだ整理が不十分な部分もある。これらの検討は今後の課題であろう。

同時に本稿では、予測市場によって専門家によるリスク判断を代替する可能性について検討してきた。予測市場は、先に述べたようにプロセスの中で関係する変数が他の予測方に比べて非常に多い。このため、実験的な環境では非常に高い精度を見せることが多いものの、実務の環境においては予測が失敗したとしてもその原因を明確に把握する検討することが難しいと言う問題を持っている。

予測は、一定のリスクを含んだものである。予測を利用するとは、その予測の中に潜在するリスクを受け入れるということに他ならない。このために予測の利用には平均的な予測力だけでなく、予測形成への理解が必要になる。将来的な予測市場の公共的利用を検討するにあたって、本稿で検討した予測市場の結果が不安定化する局面（問題状況の種類、過信の影響、運営に関わるバイアス）についての知見は予測制度設計の重要な観点となる。

註

- ¹⁾ 本稿では、予測の対象やそれを構成する要素が発信する自身の状態についてのシグナルとそのデータを「情報」、要素のシグナルやデータが示す意味を解釈し、影響を計算するものを「知識」と呼ぶ。
- ²⁾ [予測市場] アイオワ電子市場 <http://tippie.uiowa.edu/iem/>
- ³⁾ 特に選挙を対象とした得票率市場では2000年までの大統領選挙を対象におけるギャラップ社の予測の平均誤差2.1%に対してアイオワ予測市場(IEM)の平均誤差は1.4%であった。Berg et al (08)によれば、88年から04年までのIEMの964の事例の予測精度は、世論調査を元にした予測を74%で上回った。
- ⁴⁾ [政策分析市場] policy analysis market (PAM) <http://www.policyanalysis.org> (既に閉鎖)
- ⁵⁾ ヒューレット・パッカード社の事例はPlott & Chen (2002)、Google、Fordなどでの事例はCowgill & Zitzewitz (2015)で分析がある。HSXの分析はWolfers & Zitzewitz (2004)など。
- ⁶⁾ Aggregative Contingent Estimation (ACE) <http://www.iarpa.gov/index.php/research-programs/ace> [accessed 20 December 2015]
- ⁷⁾ DAは参加者が任意の時点で「売り・買い注文」を出し、システムは売り・買いの注文同士を比べ、お互いの価格が許容範囲内のものがあればマッチングさせ取引成立とする方法である。それぞれの参加者は自らの主観的信念の分布を元に、「期待利益」を最大にするように行動する。ダブルオークションでは、取引はお互いの主観的確率分布に「重複が無い」か「重複より重複しない部分が小さい」時に合意が成立する。
- ⁸⁾ 「配当を変化させる」方式(パリミュチュエル拡張(Penock04))は競馬などの勝ち馬投票券に近い。常に証券自体は一定の価格で、参加者はマーケットメーカー相手に証券の発行と買い取りをさせる。つまり予測を当てても、その証券の人気が高い(誰もが予想した)ならばその配当は小さくなり、逆に発行枚数の少ない予測証券が当たれば配当は大きくなる。参加者はダブルオークション方式とは異なり、期待配当と主観的予測を元に期待利益を計算する。
- ⁹⁾ GJP <http://goodjudgment.com/gjp/>
- ¹⁰⁾ Wintle et al (12) など
- ¹¹⁾ ノミナルグループ法はVan de Ven & Delbecq (1971, 1974)の開発したブレインストーミングを改良した手法。最初に、グループのメンバーは独立して動作し、問題に対する個々の推定値を生成。その後、グループは問題を審議するための非構造化の議論に入ります。最後に、グループのメンバーは再び独立して動作し、その最終的な個々の推定値を提供します。集団討議で発生しやすい「萎縮」を回避し、多くのアイデア(意見)を誘発する。
- ¹²⁾ Greafe (2011, 2014), Prokesch et al (2014), Sprenger (2012) など。
- ¹³⁾ 未知の事例に対して不確実性の確率評価を行う際には、時に類推を通じて「似た」事例の確率が流用される。そもそもあらゆる事例は他の事例とは同一であり得ない以上、「主観的に同一カテゴリと判断された」事例群の「確率」が流用されているといえる。

参考文献

- Atanasov, P., Rescober, P., Stone, E., Servan-Schreiber, E., Tetlock, P., Ungar, L., & Mellers, B. (2014). Distilling the wisdom of crowds: Prediction markets versus prediction polls. *Manuscript under review*.
- Berg, J. E., Nelson, F. D., & Rietz, T. A. (2008). Prediction market accuracy in the long run. *International Journal of Forecasting*, 24(2), 285-300.
- Berg, J. E., & Rietz, T. A. (2006). The Iowa electronic markets: Stylized facts and open issues. Information Markets: A New Way of Making Decisions, edited by Robert W. Hahn and Paul C. Tetlock. AEI-

- Brookings Joint Center for Regulatory Studies, 142-169.
- Burgman, M. A., McBride, M., Ashton, R., Speirs-Bridge, A., Flander, L., Wintle, B., Fidler, F., Rumpff, L., & Twardy, C. (2011). Expert status and performance. *PLoS One*, 6(7), e22998.
- Camerer, C. F., & Johnson, E. J. (1991). The process-performance paradox in expert judgment: How can experts know so much and predict so badly? In K. A. Ericsson & J. Smith (Eds.), *Towards a general theory of expertise: Prospects and limits* (pp.195-217). New York: Cambridge Press.
- Chi, M.T.H. (2006). Two approaches to the study of experts' characteristics. In: Ericsson KA, Charness N, Feltovitch PJ, Hoffman RR, editors. *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge: Cambridge University Press. pp. 21-30.
- Cowgill, B., & Zitzewitz, E. (2015). Corporate Prediction Markets: Evidence from Google, Ford, and Firm X1. *The Review of Economic Studies*, rdv014.
- Graefe, A. (2011b). Prediction market accuracy for business forecasting. *Prediction markets*, 87-95.
- Graefe, A. (2014). Accuracy of vote expectation surveys in forecasting elections. *Public Opinion Quarterly*, 78, 204-232.
- Graefe, A., & Armstrong, J. S. (2011). Comparing face-to-face meetings, nominal groups, Delphi and prediction markets on an estimation task. *International Journal of Forecasting*, 27(1), 183-195.
- Green, K. C., & Armstrong, J. S. (2007). Structured analogies for forecasting. *International Journal of Forecasting*, 23(3), 365-376.
- Hanson, R. (2003). Combinatorial information market design. *Inform.Systems Frontiers* 5(1) 107-119.
- Hanson, R. (2005). The informed press favored the policy analysis market. *George Mason University*, 15.
- Healy, P.J., Llinardi, S., Lowery, J.R., Ledyard, J.O. (2010) Prediction Markets: Alternative Mechanisms for Complex Environments with Few Traders. *Management Science*, 56(11), 1977-1996.
- Herzog, S. M., & Hertwig, R. (2011). "The wisdom of ignorant crowds: Predicting sport outcomes by mere recognition". *Judgment and Decision Making*, 6(1), 58-72.
- Kardes, F.R. (2006). When should consumers and managers trust their intuition? *Journal of Consumer Psychology* 16: 20-24.
- Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). "How social influence can undermine the wisdom of crowd effect". *Proceedings of the National Academy of Sciences*, 108(22), 9020-9025.
- Mellers, B., Stone, E., Atanasov, P., Rohrbaugh, N., Metz, S. E., Ungar, L., Bishop, M. M., Horowitz, M., Merkle, E., & Tetlock, P. (2015). The psychology of intelligence analysis: Drivers of prediction accuracy in world politics. *Journal of experimental psychology: applied*, 21(1), 1.
- McKelvey, R. D., T. Page. (1990). Public and private information: An experimental study of information pooling. *Econometrica* 58(6),1321-1339.
- McKinsey & Company (2013). Business and Web 2.0: An interactive feature, Available at: https://www.mckinseyquarterly.com/Strategy/Innovation/Business_and_Web_20_An_interactive_feature_2431 [accessed 9 July 2015]
- Nikolopoulos, K., Litsa, A., Petropoulos, F., Bougioukos, V., & Khammash, M. (2015). Relative performance of methods for forecasting special events. *Journal of Business Research*.
- Olson, K., Laskey, K., & Twardy, C. (2014). Interval Elicitation of Forecasts in a Prediction Market Reveals Lack of Anchoring "Bias". arXiv preprint arXiv:1406.7584.
- Pennock, D. M.(2004). A dynamic parimutuel market for hedging, wagering, and information aggregation,

- Proc. 5th ACM Conf. on Electric Commerce*, pp. 170-179
- Plott, C. R., & Chen, K. Y. (2002). Information aggregation mechanisms: Concept, *design and implementation for a sales forecasting problem*.
- Prokesch, T., Heiko, A., & Wohlenberg, H. (2014). Integrating prediction market and Delphi methodology into a foresight support system—Insights from an online game. *Technological Forecasting and Social Change*.
- Spence, M. T., & Brucks, M. (1997). The moderating effects of problem characteristics on experts' and novices' judgments. *Journal of marketing Research*, 233-247.
- Sprengr, T., Bolster, P., & Venkateswaran, A. (2012). Conditional prediction markets as corporate decision support systems—an experimental comparison with group deliberations. *The Journal of Prediction Markets*, 1(3), 189-208.
- Surowiecki, J. (2004). *The Wisdom of Crowds*. Anchor Books. (小高尚子訳 (2006)。『「みんなの意見」は案外正しい』。角川書店。)
- Tetlock, P. E., Mellers, B. A., Rohrbaugh, N., & Chen, E. (2014). Forecasting Tournaments Tools for Increasing Transparency and Improving the Quality of Debate. *Current Directions in Psychological Science*, 23(4), 290-295.
- Ungar, L., Mellors, B., Satopää, V., Baron, J., Tetlock, P., Ramos, J., & Swift, S. (2012). The Good Judgment Project: A Large Scale Test.
- Van de Ven, A. H. & Delbecq, A. L. (1971). Nominal versus interacting group processes for committee decision making effectiveness, *Academy of Management Journal*, 14, 203-212.
- Van de Ven, A. H. & Delbecq, A. L. (1974). The effectiveness of nominal, Delphi, and interacting group decision making processes, *Academy of Management Journal*, 17, 605-621.
- Wallsten, T. Set la (2012). The Aggregative Contingent Estimation System: Selecting, Rewarding, and Training Experts in a Wisdom of Crowds Approach to Forecasting. *In AAAI Spring Symposium: Wisdom of the Crowd*.
- Wintle, B., et al (2012). The Intelligence Game: Assessing Delphi Groups and Structured Question Formats. <http://ro.ecu.edu.au/asi/26/>
- Wolfram, T. (2015). Have Corporate Prediction Markets Had Their Heyday?. *Foresight: The International Journal of Applied Forecasting*, (37), 29-36.
- Wolfers, J., & Zitzewitz, E. (2004). Prediction markets (No. w10504). *National Bureau of Economic Research*.
- Yang, S., Li, T., & van Heck, E. (2015). Information transparency in prediction markets. *Decision Support Systems*.
- 岡野 匡志 (2015) 「"wisdom of crowd" 論について」『政策科学』23 巻1 号、53-66 頁 立命館大学政策科学会
- 岡野 匡志 (2016) 「予測市場とバイアスの所在」『創地共望』5 号 立命館大学地域情報研究センター (刊行予定)
- 宮脇 昇 (2008) 「国際政治とシミュレーション」『シミュレーション』27 巻3 号、2008 年9 月、52-54 頁。