

<研究ノート>

個人的確率の抽出に関する 1971 年の サヴェジ氏の論文の第 10 節について

園 信太郎

1. はじめに

サヴェジ基礎論で「統計学の基礎づけ」に大きく貢献したサヴェジ氏は、ほぼ 17 年後に、個人的確率の抽出に関する古典的な論文 Savage (December 1971) を公にすることとなる。サヴェジ氏自身は 1971 年 11 月 1 日に急逝しているため、この論文(つまり Elicitation (1971)) は彼の死後に公表されたのである。またこれは、Bruno de Finetti との少なくとも 10 年にわたる諸議論を背景としている。

Elicitation (1971) は、全 10 節から成り、末尾を除く(冒頭からの)九つの節では、個人的確率(及び期待値)の抽出に関する数学的な議論が展開されており、「個人」のオピニオンとしての「確率」を、でき得る限り精確に抽出するための「得点規則, scoring rule」の、定式化及び特徴づけが親切に議論されている。一方末尾の第 10 節では、表題が Practical sidelights となっているが、10.1 から 10.14 までの 14 個の小節から成り、「得点規則による「確率」の抽出」の実践に関する諸問題が議論されているのである。そこでは、「規則」と「現実」との間の隔たりに対する、サヴェジ氏の態度及び思索をうかがうことができる。なお sidelights という表現から、付足的な議論と思われそうだが、実際には「現実」に対するサヴェジ氏の真摯な思索から成り、この節だけでもほとんど独立した論述となっており、彼の(「現実」への)態度に関心を持つ者にとっては貴重な一節である。

そこでこの第 10 節をなんとか読解することによって、サヴェジ氏の思索を探索することとする。

2. 第 1 小節

この小節の表題は Scope of this section である。これは三つの段落から成り、冒頭の段落を引くと次である。

This article, which is largely about the mathematical aspects of the procedures for eliciting personal probabilities that are now often called proper scoring rules (and also admissible probability measurement procedures), would be misleading and incomplete without some discussion of their actual and potential applications. My preparation for that is inadequate; for I have done no practical work in the area nor even followed the practically oriented literature with energy and care. But it seems incumbent on me to mention, to the best of my knowledge, the sorts of applications that have been envisaged, some difficulties that threaten them, and criteria that might help in selecting among the plethora of proper scoring rules. The subject is a ramified one, and even a cursory survey demands considerable space.

ここで「適正な得点規則, proper scoring rule」と言うのは、「質問者」が「被質問者」に対して提示する（「被質問者」の利益及び「確率」に関する）「利益の分配規則」で、「被質問者」が自身の「期待される損失」をでき得る限り少なくしようと真剣に考えるのならば、その「被質問者」は、自身にとっての本来の「確率」（あるいは「期待値」）を（「質問者」に対する「答え」として）示さざるを得ない、というものである。サヴェジ氏は、自身が展開した（適正な得点規則に関する）数学的な議論が、「現実」にどのように関わるのかについて説明（さらには弁明）されることがないのならば、自身の数学的な議論は誤解されるに至るであろうと恐れているのである。なおサヴェジ氏は、自身が、「実際の応用」について門外漢であるとの態度をとってはいるが、彼は一貫して、個人的確率の応用について執拗な関心を保持し続けているのである。

なお彼はこの小節で、適正な得点規則に対して為されてきた諸実験に関する議論は、その重要性にもかかわらず、敢えて省略することをことわり、参考文献のみを挙げている。

3. オピニオンの使われかた

第2小節の表題はThe uses of opinionsである。個人的確率とは、実際の状況では、「個人」のオピニオンのことであり、そこでオピニオンの使われかたが問題となってくる。適正な得点規則によって「個人」の「確率」を本当に抽出できるとしても、我我は、「なぜ」そして「いつ」、このような個人的「オピニオン」に関心を持つに至るのであろうか。この小節の2番目の段落を引くと次である。

Often, we want to make use of the opinion of a person whom we regard as an expert. Does the weatherman think that it will rain, the doctor that we shall soon get well, the lawyer that it would be better to

settle out of court, or the geologist that there might be lots of oil at the bottom of a deep hole? Most of the following subsections are concerned explicitly with the utilization of experts.

とにかく我我は、「エキスパート, expert」のオピニオンを利用することを、しばしば欲するのである。これからの天気については、天気予報係の意見を聞きたいし、病気については、医者の見解を欲するし、訴訟沙汰では弁護士らの判断がどうであるかを知りたいし、もし石油を探索するのならば、そしてもし油井を掘削するつもりならば、地質学者などの見解を欲することであろう。

一方、「個人」がどの程度よく教育されているかを知りたい場合にも、我我はその「個人」のオピニオンを調べたくなくなるのだが、学校などで行われている「試験」は正にそのような調査のためのものと言って良いであろう。なおサヴェジ氏は最後の第14小節で強調しているのだが、得点規則が、通常の選択肢型の問題様式を、より精密なものとする可能性をきわめて重視している。

さらに、与えられている「仕事」をこなすためには、エキスパートたちの内で誰に相談すれば良いのかを、彼らが今までに提示してきたオピニオンと「現実」との関りを考慮して、各人の（その「仕事」との関りでの）「価値」を見積もることが必要となってくるであろう。なお、この小節の3番目の段落の末尾の文を引くと次である。

This domain of application shades into that of trying to determine which among possible experts are most valuable for a given task by means of the relation between their past opinions and reality.

さらに4番目の段落では世論調査が言及され

ている。public-opinion polls と言うのだから、世間のオピニオンに関するのだと思われるかもしれないが、世論調査が関わるのは世間のオピニオンではなく、主に世間の選好ではなからうかとサヴェジ氏は注意を促すのである。彼は、いまだ決着がついていない「事実」に関する世論調査が行われるとすれば、それは世間の信念に関するものであり、結局、世間のオピニオンに関わるのだと認めている。しかし、世論調査の参加者に対して、一人一人の「予測の正確さ」に応じて、いわば事後的に、「利益」を配分するというようなことは、一時的で偶然的なそれらの参加者を対象とするのならば、実際上不可能ではなからうかと、彼は判断しているのである。

さらに、この段落の末尾の二文を引くと次である。

Yet scoring rules might be used in training panelists to assess their own probabilities, perhaps partly in paid practice sessions about immediately verifiable predictions. This would seem less farfetched should scoring rules come to be widely applied in the schools.

これは得点規則を、ゲームや討論などの回答者や討論者が、自身の「確率」を評価するために利用する状況に言及しているものであり、このような利用を通して、自身の本来のオピニオンを「知る」力が徐々に養成されるはずだと言うのである。参加者は、参加することから（現状よりも）損をすることはなく、しかし一方、すぐにも当否が示されることとなるであろうような、予測を為すのであり、自身の予測と実際の結果とに応じて、但し、問題の得点規則に従って、各自は、利益あるいは損失を受け取ることとなるのである。なおサヴェジ氏は、得点規則が広く学校教育の現場で利用されることとなれば、このような「得点」の配分は不自然ではな

くなるであろうと、期待している。

この小節10.2の末尾の段落は実にサヴェジ氏らしい文章なので、これを次に引く。

One domain of potential application of proper scoring rules is suggested not so much by the problem of obtaining the opinions of others as by the difficulty of obtaining our own. For it is by no means easy to elicit your own probabilities. Vagueness is a major obstacle, and your first reactions are often greatly modified when you reflect upon their implications. Those who have experimented on themselves and on others generally feel that frequent practice with proper scoring rules and with other probability elicitors helps a person to combat vagueness and to arrive more promptly and accurately at his personal probability. This proposition might be difficult to investigate experimentally and even seems difficult to state with precision, but its promise of benefit is great. (See, for example, [12,59].)

つまり、適正な得点規則の潜在的な「真の力」は、他人ではなく正に自分自身のオピニオンを獲得しようと試みる際に、正に自分自身に対して（暗黙の内に）「明らか」となる、というのである。サヴェジ氏はここで一人一人の「己証」に訴えているのである。自分自身の「確率」をつきとめるために、適正な得点規則や他の抽出法を自身に当てはめて行く過程において、自身のオピニオンに（不可避的に）まといつく「漠然性、vagueness」と戦う態度が徐々に練成され、より速くかつより正確に「自身にとっての本来の確率」に到達できるようになることが、自身の「一人（いちにん）の」実践の場で「実感」されるようになるというのである。しかし、このような過程を定式化し、この「実感」の存在を証明することには、独特の

困難が伴う。何しろ個人的確率の立場からすれば、「確率」とは、「最終的には」一人一人が「個人的に」定めなければならない「存在」であり、(未知ではあるが固定されている)「真の確率」などは、(ぎりぎりの所では)元来「存在しない」からなのである。つまり自身が(論理や経験を徹底的に援用して)定めた「確率」が「本来の確率」なのであり、他に「真の確率」を求めることなど「在り得ない」わけである。サヴェジ氏は、得点法則の「真の力」は、一人一人の「この己証の場」において明らかになると視ている。

4. 第3小節

これは Yes, No, or Maybe is not enough という表題である。サヴェジ氏は、競馬などの「かけ」に関するエキスパートを除いては、エキスパートが(自身の)オピニオンを「確率」で表すことは、ごく最近まで、あまり為されてはいなかったと、注意を促している。彼は、医者の診断や、地質学者による(油井を掘るための)場所の推測、そしてようやく「確率」を使い始めている天気予報などに言及している。二番目の段落を引くと次である。

The importance of a system of communication in which experts express themselves in terms of genuine personal probabilities and in which those who utilize the opinions of experts—that is, all of us—are trained in understanding and using such probabilities was energetically underlined by Grayson [27]. In recent years, meteorologists have been announcing forecasts in terms of probability on some radio and television services, but not all who offer these probability forecasts think in terms of personal probability, and some seem to be very vague indeed about what they mean by probability in a forecast. The earliest known refer-

ence to proper scoring rules is by the meteorological statistician, Brier [5], and much of the current literature on proper scoring rules is inspired by meteorology, as in [21, 39, 40, 41, 46, 53, 58, 59, 60] and works cited in them.

サヴェジ氏は、エキスパートらが「正真正銘の個人的確率」によって、自身のオピニオンを表明することを望んでいるのであり、エキスパートの「確率」を利用する我我も、「本来の確率」とはいかなるものであるのかを理解し、かつ「その確率」を利用するための、正しいトレーニングを受ける必要があることを指摘しているのである。しかし、気象予報士が利用するようになってきた(予報の)「確率」を、予報士自身が良く理解しているとは限らないのであり、実際、あまりにも漠然とした意味合いで「確率」という言葉を利用している場合もあるのである。なお彼は、「適正な得点規則」の開発や利用についての地道な議論が、気象に関する統計家らによって行われてきた事実注意到を促している。

例えば、「明日の朝6時から昼の12時までの間に某地において雨が降る確率」という事柄を、「(予報をしている)現在の状況と「同様の」状況が多数回繰り返されるのならば、「問題の雨が降る相対的な頻度」と見なした場合、その際の「確率」とはかなり漠然としたものであり、実際、「「同様の」状況」という場合の「同様の」とは、いかなる判断に基づくのか、また、「多数回繰り返す」とはどのようなことなのか、さらにはなぜ「相対的な頻度」なのかなど、明晰でない点があるのであり、しかも問題となっているのは、「繰り返し」ではなく、「明日」の特定の時間帯の「一回限り」の事象(即ち「できごと」)なのである。やはり予報士は、「自身の確率」を述べなければ「ならない」はずである。

この後サヴェジ氏は、適正な得点規則と現実

のエキスパートとの関りについて、注意を促している。つまり、適正な得点規則そのものは、失点をでき得る限り避けようとする営みを通して、エキスパート自身が、「確率」とはいかなる事柄なのかを理解し、さらには、自身が報告している値が本当に自身の「確率」なのかを自問するための、「緩やかな刺激, a mild incentive」を(エキスパート自身に)与える、一つの器具と見なし得るのである。だが一方では、石油探査に関心のある地質学者が、ある油井の掘削に関っている企業の株を買うとすると、この「株を買う」という行いは、「問題の掘削作業に伴う(その企業の)期待されるコストよりも、その油井がもたらすであろう(その企業の)収益の方が大である」との、その地質学者の(問題の油井に関する)オピニオンを、明白に指し示しているのである。この場合、「株を買うか否か」という判断は、ある種の得点法則の遂行に対応すると見なし得るのである。つまり、少なくともその地質学者にとっては、「その掘削の結果石油は発見され、油井は成立し、しかも採算が取れる」と見込まれるのである。従って、「その地下には十分な石油がある」と(エキスパートとしての)その地質学者は判断していることとなる。

この「油井の例」に続く段落、つまりこの小節の末尾の段落を引くと次である。

This same example, which suggests so vividly how any proper scoring rule might be applied to elicit from a geologist his opinions about the yield of a well that is to be drilled, also serves to illustrate what seems to be an insurmountable obstacle to the application of arbitrary scoring rules, and even of any proper scoring rules to all, to many situations in which opinion is sought as a basis for decision. If whether the well is to be drilled does not depend on this geologist's opinion, then any proper

scoring rule concerning its expected yield can be implemented, in particular by offering to sell shares in the well at various prices. But if whether to drill the well depends in part on the opinion of the geologist, then some of the events about which his opinion is wanted may never be tested precisely because of his advice, so the scoring-rule contract cannot be offered literally. The phenomenon is ubiquitous. We can never know what would have happened had surgery been ventured, had a certain product been marketed, or had a certain student been admitted. Business sharing, to be discussed in Subsection 10.12, seems to offer some possibility of circumventing this widespread obstacle to the literal application of proper scoring rules.

ここでは、得点規則を実際に利用する際の困難が言及されている。油井を掘削するか否かの判断が、「確率」や「期待値」の抽出の対象となっている)その地質学者のオピニオンに依存しない様式で行われる場合には、適正な得点規則を(得られるかもしれない石油の配分を、いろいろな価格で地質学者に申し出ることなどを通して)遂行することができるであろう。しかし、油井を掘削するか否かの判断が(部分的にはあっても)その地質学者のオピニオンに依存するのならば、彼のオピニオンが求められている事象の内の幾つかは、そのオピニオンの適否の程度がためされないまま残されるのである。例えば、その地質学者が、「その場所には採算が取れる程の石油はない」と判断した場合、企業側は、「掘削はしない」と判断するものとしよう。すると、彼が実際に「石油はない」と判断すれば、彼のオピニオンの適否はその企業側によってためされることはなくなるのである。「もしこれこれしたとすれば、どうなっていたであろうか」と問うことはできても、その答えは、

知られることはないのである。

5. 大きな金銭的損得

第4小節の表題は Big money である。冒頭の段落を次に引く。

The interrogator has an interest in making the possible fluctuations in the wealth of the expert large. For this motivates the expert to reflect hard and well before answering and yet need not add systematically to his expected fee. The larger the fluctuations are, however, the more the expert is motivated to report not his real probabilities but numbers that reflect in part the nonlinearity of his utility. The theory of these distortions has been somewhat explored by Winkler and Murphy [60], and some possible ways to avoid them or compensate for them were reported on in Section 3. My own hope and expectation is that the skillful use of small, or even purely symbolic, scoring-rule payments (in addition to the usual compensations) will enable experts to know and to communicate their opinions much more accurately than has been usual.

(エキスパートからオピニオンを抽出しようとする) 質問者は、(問答に参加することによって) エキスパートが損害を被ることがないように、少なくとも金銭的な配慮をする必要性を感じることであろう。このような言わば問答参加料は、エキスパートが問答への参加を受け入れてくれるための、実際上の必要条件である。だがさらに質問者は、(エキスパートの参加のみでなく) 彼が「真剣に」質問に答えてくれることを願うはずである。質問者はできれば彼が、一連の質問に対して、熟慮してくれることを望む。そこで質問者は、(彼に対して提示される)

得点規則が彼にもたらす(少なくとも金銭的な) 損得の幅を大きくすることを、企図するかもしれない。しかし、金銭的損得の幅を大きくすれば、彼は、自身の「効用」の非線形性を反映した言わば「歪んだ」値を告げることとなり、彼自身の現実の「確率」は(少なくとも質問者にとっては) 隠されたままとなるのである。そこで、金銭的損得の幅を直接に拡大するのではなく、小額の支払いのみを利用した、さらには全く象徴的な模造貨幣のようなものを用いて、「効用」の非線形性がもたらす「歪み」を迂回して、エキスパート自身の「確率」を抽出するやりかたが強く望まれるのである。

この第4小節はこれに続く段落で終了しているが、サヴェジ氏はそこで、Ward Edwards が「[かけ] をするという行動」について、興味深い観察を行っていることを注意し、さらに、得点規則を適用する際の支払いに、本物の通貨や、私的な代用硬貨を利用する場合でも、法律や経営管理に関する問題を惹き起しかねないであろうことを、認めている。

6. 微小な確率

第5番目の小節は、Tiny probabilities and the expert であり、二つの段落から成り立っている。冒頭の段落を引くと次である。

Probabilities corresponding to odds of one in a million or even less, such as the probabilities of specific disasters are sometimes important [32]. Any scheme to give an expert a serious cash incentive to reveal his personal probabilities for very improbable events would seem to court insuperable difficulties with the nonlinearity of utility. And, even if we do not despair of obtaining sincere opinions about such matters by engaging sincerely inclined experts in the right kind of make-believe, the difficulty of such make-believe, and the special training

required for it, are in danger of being greatly underrated.

問題となっている大災害が現実のものとなる「確率」というような場合の「その確率」は、状況によっては、「極めて」小さなものとなりうる。「その確率」を誠実に見積もるためのエキスパートが存在するとしても、彼自身にとっての「その値」を定めることは、極めて困難であり、彼自身の「効用」の非線形性が「その値」を歪める恐れがあるのである。ここで用心すべきことは、「その値」を定める作業は、言わば「空想にふけること」だということである。この種の「空想に「真剣に」ふける」専門家を養成することは容易なことではない。

サヴェジ氏はこれに続く段落で、いかにも彼らしい例に言及しているのです、これを引く。

Magnification is sometimes possible. For example, though it would not be practical to engage me in a meaningful bet that the birth awaited by the recently married Smiths will be quadruplets, my expectation for the number of sets of quadruplets among the next million American births (about two, based on a little reading) could be elicited and might be of some use. But this relief is largely illusory. For a real expert on multiple births would be expected to take into account such data as that Mrs. Smith is herself a twin and a very young bride, so there is no practical possibility of counting the quadruplet births in a large number of cases similar in the respects considered pertinent.

「微少な確率」を定める際の困難を彼は例示する。近頃結婚したSmith夫妻の間に子供が生まれることとなっているのだが、この出産が四つ子をもたらす「確率」が問題となっている。

この「確率」を評価するために、実際的とは思われないのだが、とにかくサヴェジ氏が雇われたとする。そこで彼は、読書で得た自分自身の知識から、米国での（これからの）百万回の出産の内で四つ子となるのは約二回であるという比率を、提示したとしよう。つまり「約百万分の二」という「確率」が（多胎出産の専門家ではない）サヴェジ氏から抽出されたわけである。「この値」は、それなりの利用の仕方があるのかもしれない。しかし、「この値」が何らかの安堵感をもたらすとすれば、その多くは幻想である。多胎出産の本物のエキスパートであれば、例えば、Smith夫人自身は双子の内の一であり、まだ極めて若い、とか言うようなデータを考慮に入れることであろうし、このようにして問題の「確率」に関りがあると思われる観点を取り上げていくと、「それらの観点が類似している多くの場合における、四つ子の出産回数を求める」という作業が実際上不可能となるのである。つまり、本物のエキスパートは、最終的には、自身の個人的確率を提示せざるを得ないであろうが、しかしそれは相当に困難なのである。なおサヴェジ氏は、「事例の比率」と（本物のエキスパートの）「確率」との間の葛藤を十分に承知しており、この葛藤との関りで、次の小節で perfect fool という言葉を発している。

7. perfect fool と言う言葉

第6小節は Employing expert opinion である。「専門家の意見」をいかにして利用するのが（我々にとって）得策なのか。この小節は二つの段落から成り、冒頭の段落を引くと次である。

In just what ways can you expect to profit from the opinions of experts in serious matters? According to a very broad model, you have an important decision to make in the light of all sorts of data at

your disposal, and this data may include the behavior of experts. You could, in principle, explore all sorts of ways of interrogating an expert—not confining yourself to eliciting his opinions—and study empirically how his responses can profit your business. Conceivably, pain in the weatherman's great toe would better help you plan picnics than would his opinion about the weather. Yet I presume that ordinarily little of importance would be lost if you could obtain only the opinion of the expert, that is, his personal probabilities. What should be done with such an opinion? The simplest thing, and sometimes the appropriate thing, would be to make the expert's opinion your own.

「決定」という事柄を真剣に考え抜くのならば、専門家の振舞いのみでなく、利用可能なあらゆる「データ」を利用して、自身にとって重要である「一つの決定」を下すべきなのである。専門家の意見のみでなく、その行動のあらゆる側面が問題となるのであり、それらがもたらした自身への影響が（自身にとっての）利益との関りで慎重に考察されなければならないはずである。こちらの質問に対する（その専門家の）反応が、自身の仕事にどのような利益をもたらしたのかを経験的に調べることは重要なことであり、さらには、例えば、「その天気予報係」の足の親指の痛さが、（自身の）ピクニックの計画にとって、「その天気予報係」の）天気に関する意見よりも、重要なデータともなり得るのである。しかしサヴェジ氏は、通常は、エキスパートのオピニオンが問題なのであり、結局彼の個人的確率のみが（「決定」を為す者にとって）重要なのだと、判断している。このエキスパートのオピニオンに対する利用の様式は、最も単純なやりかたは、彼のそのオピニオンを、自分のものとして採用してしまうことなのである。

だが、この単純なやりかたが常に正当であるとは限らない。サヴェジ氏は、これに続く段落の二番目の文で perfect fool という表現を使っているので、（この問題の段落を）次に引く。

This is by no means mandatory. For example, an expert who always ascribes very small probability to what actually occurs and to nothing else would be as useful as one who is omniscient, but you would of course not make the opinions of this perfect fool your own. Again, you might discover with experience that your expert is optimistic or pessimistic in some respect and therefore temper his judgments. Should he suspect you of this, however, you and he may well be on the escalator to perdition.

つまり、「実際に起っている事柄に対して非常に小さな確率を配分し、他の事柄に対してはそのような配分はしない」という態度を常に取り続けるる専門家は、本来のエキスパートではなく、「物知り」の類とかかわらないというのである。このような百科事典的な固着状態をいつまでも取り続ける専門家を、サヴェジ氏は perfect fool と呼び、このような専門家の意見を鵜呑みにする必要などないと言うのである。さらに自身が相談している専門家は、経験的に見て、ある種の局面については楽観あるいは悲観的な態度をとる傾向があると、「あなた」が判断する場合には、「あなた」はその専門家の意見を「加減する」かもしれないと、彼は指摘するのである。だが彼が指摘するように、「あなた」が「加減する」かもしれないと「その専門家」が疑うのならば、「その専門家」はあらかじめ自身のオピニオンを「加減する」ことであろうし、次には「あなた」がこの「疑念」及び「加減」を疑えば、さらなる「加減」を、「あなた」は試みるであろう、というように、「あなた」

と「その専門家」とは、極めて悲惨な状況へと墮ちて行くのである。

8. オピニオンの隔たりをどう処理するか

第7小節は Divergent opinion であり、ただ一つの段落から成る。これを次に引く。

We often have access to more than one expert, and what to do when doctors disagree has always been, and will always be, a quandary. One important thing to do, but far outside the scope of this paper, is to encourage the right kind of communication between the experts. Exploration of how to do this is, for example, one of the aims of the Delphi technique [3; 8; 29, Part II]. In general, good communication is what makes the experts share factual information and help each other think their opinions through thoroughly, and bad communication is what encourages various vices such as exaggeration and excessive deference. Sooner or later, despite all techniques of communication, divergent expert opinions will have to be faced. Perhaps you will make some composite of one or more expert opinions and your own opinion. An extreme way to do that would be to decide, on the basis of past experience or otherwise, that a particular one of the experts is the only one worth listening to and to make his opinion your own. A more general procedure would be to average the opinions, that is, to average the probability distributions associated with the experts (possibly including yourself), giving each the weight you think appropriate. Thus, rather than simply choose one expert among several, you can choose among the infinite number of synthetic experts that constitute the convex closure of

the several. (A few key references bearing on the subject of this paragraph are [10; 11; 47, Ch.10; 52, p.65 ff; 56].)

ただ一人だけのエキスパートのオピニオンのみに従うという流儀もあるかもしれないが、複数のエキスパートに相談することも当然起り得る。しかし、エキスパートの間で見解が異なり、さらには対立するかもしれない。このような異議や対立の程度を軽減するために、エキスパート間の意思疎通を円滑に行い、「事実」とされる事柄を各自が共有し、意見の誇張や甚だしい服従をでき得る限り消去して行くことが多分可能である。だが、不確定性が存在する限り、基本的な「見解の対立」とも言うべき事柄が、多分残されるであろう。ここでも、ただ一人のエキスパートのみを選んで、その見解に従うという流儀は採り得る。しかし、オピニオンは結局個人的確率で表現されるとするサヴェジ氏の見解によれば、オピニオンの凸結合もまたオピニオンと解釈し得る。つまり、複数の見解の凸結合の全体から一つの（合成的な）オピニオンを選択すべきだとの立場が意味を持ち得るのである。ここで注意すべきなのは、凸結合を生成する際の重みづけは各自が「主観的に」為すということである。サヴェジ氏は、あくまでも「決定の主観性」を譲らないのである。なお、彼が明確に言及しているように、重みづけの対象として「自分自身」を含めることも可能である。また、この「主観的な重みづけ」を、彼は average という動詞で表現しているが、これは「過去の経験」などに基づく、（自分自身をも含み得る）エキスパートへの、「主観的な値踏み」としての「重さ」を用いた、「平均」である。ここで用心すべきなのは、「過去の経験」から（少なくとも原理上は）「客観的に」定まるであろう「重さ」ではなく、「過去の経験」などから「主観的に」定められる「重さ」ということである。

9. 道具としてのエキスパート

第8小節は The expert as an instrument である。冒頭の段落を引くと次である。

When is one expert, real or synthetic, to be preferred to another? An "expert" in this context is a mechanism, possibly with human components, generating numbers that you contemplate instead of your own personal probabilities (or, more generally, expectations) in certain contexts. One crude, practical answer sometimes available and appropriate, is this: Employ, until you have further experience, that expert whose past opinions, applied to your affairs, would have yielded you the largest average income.

ここでは「エキスパート」という言葉が、個人的確率の立場から一般的に捕えられており、それは、「あなた」が（自分自身の個人的確率のかわりに）利用しても良からうと判断する「数の系列」を生成する、「機構, mechanism」のことなのである。ここで、「個人的確率」に限らず「個人的期待値」としてもかまわないし、「機構」と表現しても、「人」や「人的な要素」が組み込まれていて当然良い。また、「合成的な, synthetic」エキスパートとは、前節で述べた「重みづけ」によって形成されるエキスパートである。そこで、どのような状況で、「一方のエキスパートを選ぶことが他方を選ぶことよりも得策である」と判断されるのが問われることとなってくる。この「問い」に対する、一見すると実践的なように思ってしまうかもしれない「答え」とは、「あくまでも仮にだが、「彼」の過去のオピニオンたちを「あなた」の（当時の）仕事にあてはめたとしてみよう。このような仮定法的な（想像上の）実験を行って見て、最も大きな収入を（「あなた」へと）もたらしたであろう「その彼」のオピニオンを、採用すればよい。」というものである。この「常

識的な答え」へのサヴェジ氏の態度を見るために、これに続く段落を引く。

No rule of this sort can claim absolute or objective validity, and this one has been couched especially roughly for the sake of simplicity. For example, actual past experience with the experts may be extensive, moderate, meager, or absolutely lacking. When past experience is extensive, but not too extensive, the rule often has much to recommend it; when direct past experience with the experts is meager, the rule is silly; and when such experience is altogether lacking, the rule results in a tie and is therefore empty. Actually, if you have little or no past experience with the experts, you will have to ponder them in terms of whatever information it was that brought you to regard them as promising in the first place; this well finder is regarded by the whole neighborhood as infallible with the hazel fork; that one is a professor of geology and the author of an important treatise on sub-surface hydrology but has never before tried to help anyone locate a well. In such a context, the subjective aspect of your decision is thrown into prominence, but no matter how much direct past experience you may have with the experts, the ultimate subjectivity of your choice among them never disappears, though its effects may become less agonizing—according to the personalistic Bayesian theory of statistics, as in [48].

彼は、「常識的な流儀」が持ち得る有用性をひとまず認めはするのだが、その根拠は定かではないとしているのである。この「流儀」に（合理的で）客観的な根拠があるとは言い難い。

「あなた」が問題としているエキスパートたちについて、「あなた」が多くの（しかし多すぎない）「過去の経験」を保持しているのならば、この「流儀」は（推薦に値する程に）有用であろう。しかし（それらエキスパートらについての）「直接的な経験」が貧弱であれば、この「流儀」の説得力は、かなり怪しいものとなる。さらには（「あなた」にとって）「使える経験」が全く欠如しているのならば、この「流儀」は無内容であり、それらエキスパートらは事実上対等である。仮に（「あなた」が問題としている）エキスパートらについての（「あなた」の）「経験」が、ほとんどないとしてみよう。すると、一体いかなる「情報」との関りで、「そのエキスパート」を有望であると（初めに）「あなた」が思うに至ったのかを、「あなた」が自身に問うことが重要となってくる。この「情報」とは、例えば、「彼（あるいは彼女）は、近所の皆から、井戸掘りの名人であると見なされている」とか、「彼は地質学の教授であり、地表下の水文学について重要な論文を書いているが、しかし、実際の井戸掘りを手助けしたことは一度もない」とか、「彼はどうも自分と相性が良い様だ」とかいうように、とにかく多様であるだろう。このような場合、「決定」と呼ばれる事柄が持っている「あなた」にとっての主観性は、極めて顕著なものとなる。だが、「あなた」が、「直接的な過去の経験」をいかに多く保持していようとも、「決定」と呼ばれる事柄が持つ「あなた」にとっての主観性が、消え去ることなどないのである。この「主観性」は、「あなた」に決断を迫るのだが、それ故に、「あなた」に苦痛をもたらすことであろう。しかも、個人的確率の理論によれば、いかに「情報」が豊かであっても、「苦痛」が軽減されることはあっても、ここで問題としている「決定」の究極的な主観性は存続するのである。

さらに、これに続く（二文から成る）段落を引く。

When your past experience with the experts is very extensive indeed, it may become profitable for you to refine the original rough rule by dividing up circumstances into categories and confiding in different experts for different categories of decisions. Such discussion could continue indefinitely; for the situations are innumerable and tend to parallel the whole field of decision-theoretic statistics.

「それらエキスパートら」に関する「あなた」の経験がただ単に「多い」のではなく、「極めて多い」と表現する方が多分適切であるような状況では、「あなた」の決定が求められている状況を幾つかの「類」へと分割し、それぞれの「類」に対して、「あなた」が最善であると判断するエキスパートに（心を開いて、じっくりと）相談するというやり方が考えられる。この流儀では、「最善の」エキスパートは「類」が異なれば異なり得る。また、「過去の経験」が「極めて」多いので、（類似する状況からなる）「類」に対応する「あなた」の経験は、「少なすぎる」ということはない。「経験」が増大するに従って、より細分された「類」たちが設定され、このような分岐は、「際限がない」と表現したくなる程に続くであろう。なおサヴェジ氏は、この段落に続く（唯一の文から成る、この第8小節の末尾の）段落で、PIP（Probability Information Processing）に言及している。

10. つかみにくい「効用」

第9小節は、Slippery utilitiesである。冒頭の段落を引くと次である。

A different kind of complication in applying, or adapting, the rough rule is that the notion of average income may not be read-

ily applicable to your affairs. For one thing, it may be important to measure your own income in utility rather than in cash. If this involves only determining your utility for cash, you may be able to do that reasonably well with moderate effort. If, however, the consequences of your act are not easily converted into cash but involve values difficult to weigh against each other such as beauty, justice, and health, your dilemma may be especially severe. Raiffa [44] has recently published a book largely on these subtle problems of pondering the imponderable and evaluating the invaluable that reviews, and contributes to, a considerable literature.

上の議論で言及した（エキスパートらの選択に関する）「常識的な流儀」には、さらに込み入った事情がからんでくる。「あなた」は、「そのエキスパートが（「あなた」へと）もたらずであろう「平均的な収入」を問題とするのだが、「平均的な収入」という概念がうまく当てはまらない状況が起り得る。実際、「現金」ではなしに（「あなた」にとっての）「効用」で、「収入」を測った方が適切であるかもしれない。この場合、「貨幣の「あなた」にとっての効用」のみが問題ならば、いわゆる「貨幣の効用関数」を（「あなた」の立場から）導入することで、難所を切り抜けることができるであろう。だが、「あなた」が自身の選択の結果として荷うこととなるであろう自身の「収入」は、容易には貨幣へと変換できないかもしれない。問題の「収入」には、美、正義、健康、真理、善などと呼ばれる「価値」たちがからんで来るかもしれないし、この場合「あなた」は、これらの「価値」を比較し考量しなければ「ならない」状況へと追い込まれるのである。「測り難いもの」を「測り」、「評価し難いもの」を「評価する」ことは、「あなた」を窮地へと導く。

サヴェジ氏は、この窮地への対応策として Raiffa (1968) に言及しており、さらなる考察を読者にまかせている。

これに続く段落を引く。

Another difficulty in measuring utility is seen in this example. You are the person responsible for choosing which of several televised meteorological forecasters shall serve your city. There is, I assume for simplicity, abundant evidence of past performance, and the members of the community who use the forecasts will behave in accordance with the probabilities announced in them. Since yours is a public trust, you would like to choose the forecaster that in the past would have maximized the mean income of members of the community.

「効用」の「測り難さ」の指摘がさらに続く。ここでは天気予報係の例が言及されている。仮に「あなた」が天気予報係を選ぶという「公的な責任」を負う立場にあるものとする。問題の天気予報は、「あなた」が住んでいる地域のものであり、テレビで放送され、候補となっている予報者らの過去の実績については充分量の証拠があり、しかも問題の地域の住民たちは、その予報係の告げる「確率」を自身のものとして、自らの行動を定めるものとする。（公的な地位にある）「あなた」の「取り分」は、いわば「公的な信頼」である。すると「あなた」は、「その予報者の過去の予報の悉くを仮に放送したとすれば、問題の地域の住民の全体に対する「平均的な収入」を最大化したであろうような、その予報者を天気予報係として採用する。」という流儀で、自身が置かれている状況を切り抜けようとするかもしれない。だが、ここでの「平均的な収入」とはいかなるものであろうか。

さらにこれに続く（この小節の末尾の）段落

を引く。

Subtle welfare-economic decisions about the relative importance of bent-pin anglers and barn painters could complicate your problem, but, even more important, relatively little is really known about the uses to which public weather information is put and what its economic consequences are. Thought has been given to the problem of the economic value of meteorological forecasts, both for the general public and for special purposes, but difficult, important empirical aspects of the question remain to be explored. (Key references are [36, 40].)

いったい「天気に関する公的な情報」はいかなる事柄のために利用され、それらの「利用」によってもたらされるであろう（公的な、あるいは私的な）「経済的な結果の総体」とは、いかなるものであろうか。

11. どの得点規則か？

第10小節は Which scoring rule for the trained respondent? である。「質問者」の「質問」に対して「応答者」は「数値」を答えるのだが、「質問者」はあらかじめ「応答者」に対して得点規則を提示しており、この得点規則と（答えとしての）「数値」、及び問題となっている「事象」の生起及び非生起とにより、「応答者」の「取り分」が決まるのである。ここで、「事象」が「事象の列」の場合は、（答えとしての）「数値」は「数値の列」となる。ここで「応答者」が、「自身の期待される取り分」を最大化しようとするれば、「（答えとしての）「数値」あるいは「数値の列」として、自身の個人的確率を答えざるを得ない」のならば、その場合には、「質問者」は、「適正な得点規則, proper scoring rule」を提示していると表現されるの

である。自身の「確率」を探索するという作業についてとにかく訓練をつんだ「応答者」が、trained respondent であり、彼（あるいは彼女）に対していかなる「適正な得点規則」を提示すべきなのかというのが、この小節での「問い」である。

冒頭の段落を引く。

There are as many proper scoring rules for a trichotomy, for example, as there are convex functions over the barycentric triangle, or two dimensional simplex. It would therefore seem important to study in what respects one scoring rule is better than another. But this question has thus far proved surprisingly unproductive. Its elusiveness is brought out by the consideration that an ideal subject responds to all proper scoring rules, including those involving extremely small payments, in exactly the same way. Therefore, any criteria for distinguishing among scoring rules must arise out of departures of actual subjects from the ideal.

問題となる事象（あるいは事象列）が特定されていても、「適正な得点規則」は無数に存在しているのである。つまり、応答者（あるいは「被験者」）から（問題となっている事象に対応する）「確率」を抽出するための「得点規則」は、無数にある。そこで、どの「適正な得点規則」が「その」応答者にとって好都合なのか、つまりより「精確に」問題の「確率」を抽出できるのかが、問題となってくる。だが一方、「理想的な」応答者は、どの「適正な得点規則」に対しても、常に自身の個人的確率を端的に答えるのみで、彼には「得点規則」の間の（自身にとっての）適不適など元来存在しないのである。ここで「理想的」とは、サヴェジ氏の意味で「合理的」ということで、特に、「理想的な」個人は、どのような状況においても計算上の過

ちを犯さないのである。「得点規則」の（「その」）応答者にとっての）適不適の問題は、いかに訓練を積んだ応答者であれ、とにかく「理想的な」応答者からは大きく離れた地点にいるという、「現実」に立脚した「問い」である。例えば、「現実の」応答者は極端に小額の「取り分」に対しては、真剣に答えようとはしないであろう。

これに続く段落を引く。

Since we all do depart markedly from the ideal, it might seem that one proper scoring rule would be much more effective with a real person than another, and this presumably is often so. But if a person is reasonably sophisticated, though far from ideal, the form of a proper scoring rule for eliciting his probabilities, for, say, a trichotomy should—provided its amplitude is sufficient to command attention—have little or no effect on his response. To see this, put yourself in his place. You are offered a contract that will result in certain cash payments to you depending on your choice of three numbers p , q , and r and on whether a certain game ends in win, lose, or tie. If you know what personal probabilities are and understand that the contract is so drawn that it is to your interest to report your personal probabilities, then the details of the contract seem unimportant; for no matter which proper scoring rule it corresponds to, you should ask yourself what your personal probabilities for the three events are and report them.

そこで一般的に、「現実の」応答者にとって（一般的な傾向として）より効果的な「得点規則」というものがあるであろうか、という「問い」が出てくるのである。恐らくはこの「問い」

の答えは、実際上は、「ある」である。だが「確率」という事柄を掘り下げてみる必要がある。個人的確率に関して知的な（但し、程程で良い）訓練を受けた「個人」は、「適正な得点規則」というものが、自身が（自身にとっての）「確率」を、即ち、（自身にとっての）個人的確率を、でき得る限り精確に告げることによって、また、その場合に限って、（自身にとっての）最大の期待利得をもたらすことを理解するであろう。このような「個人」にとっては、提示される「得点規則」が「適正」であることのみが大事なのであり、他の細目は問題にならないのである。「その個人」は、自分自身に対して、（問題となっている事象あるいは事象列に対する）自身の個人的確率とはいかなるものかを、（行動的に）尋問すれば良いのであり、そして「その値」を彼は、質問者（あるいは尋問者）に対して告げるのである。ひとたび数学的諸形式によって表現される個人的合理性を了解してしまえば、他の細目は末節となる。サヴェジ氏は、数学的諸形式の持つ力を良く視ている。

さらにこれに続く段落を引く。

Yet, the terms of the contract might make a modest difference to you. Suppose, for example, that very little money is to change hands in case of a tie, no matter what your response is. In this case, you have little incentive to ask yourself carefully the probability of a tie and are thus left free to focus on the relative probability of a win given that there is not a tie. In this case, your questioner will be well served if he is mainly interested in that conditional probability, and he will be badly served if he particularly wants to know the probability that you attach to a tie.

だがさらに論点は続く。仮に「その試合」がこちらの「勝ち」か、「負け」か、それとも「引

き分け」かが、問題となっているでしょう。提示されている「得点規則」によると、「あなた」が冷静に答える限りにおいて、「引き分け」の場合には、「あなた」にとってほとんどないに等しい金銭しか動かないことが明白であれば、「あなた」は、「引き分け」の「確率」(r)とは真剣に取り組まずに、「引き分け」がないという条件のもとでの、「勝ち」か「負け」かの相対的な「確率」($p : q$)を見積る作業に（勝手に、明白に意識しているか否かは別として）とりかかるかもしれない。質問者が、「勝ち」か「負け」かの（「あなた」にとっての）「条件つき確率」を強く欲しているのならば、この状況は（質問者にとって）悪くない。しかし、「引き分け」の「確率」を特に欲している場合には、この状況は良くないのである。このような（質問者にとっての）危険は、（「応答者」である）「あなた」が個人的確率の基本に通じているとしても、多分充分に起り得ることである。個人的確率を良く理解しているのだと承知している場合でも、人である限り、ほとんど意識せずに、人間的で心理的な「過ち」を充分に犯しうるのである。

さらにまたこれに続く段落を引く。

Thus, at least a vague criterion applicable even to sophisticated but human respondents emerges. Insofar as responding is hard work, the scoring rule should encourage the respondent to work hardest at what the questioner most wants to know. If this is to be effective, it must not merely be mathematically true but also plain to the respondent that he will be rewarded most for working on the right aspects of his opinion. This is in part a psychological question of human communication, subject to much speculation and experimentation. To illustrate, you are faintly curious to know whether the respondent thinks that a

tie is likely and desperate to know whether he believes that the home team will win. You might be best served by entering into two palpably separate contracts with the respondent, a small one hinging only on whether there is a tie and resulting in a rough casual elicitation of the respondent's r , and a larger one, involving no payment in case of a tie, resulting in a well considered evaluation of the ratio p to q for the respondent. Of course the two contracts together amount to a single scoring rule, though presenting them separately might work better psychologically.

自身の個人的確率の抽出において経験を積んだ応答者であっても、実際には人であるので、（質問者が提示する「得点規則」の様式によって誘導される）自身の心理的な状況から、全く影響を受けないというわけではあるまい。例えば「あなた」が、「この試合」が「引き分け」となる「確率」を「応答者」がどの程度に見積るのかについて、あまり関心がなく、一方で、「この試合」が「勝ち」となる「確率」を彼がどの程度に見積るのかについては、ぜひとも知りたいと思っているのならば、明白に分離されている二つの「得点規則」を、彼へと提示することが得策であるかもしれない。つまり、一つは、「引き分け」か否かにのみ依存する「小額」しか動かない得点規則」であり、他方は、「勝ち」か「負け」かに依存して「かなりの額」が動き、「引き分け」の場合には何の支払いも為されない、得点規則」である。個人的確率の理論のみを尊重するのならば、これら二つを一つの「得点規則」へと結合しても何らかまわらないのだが、「あなた」が提示する「規則」が（応答者に対して）持つ心理的な側面を無視し得ないと、「あなた」が本気で判断するのならば、「あなた」は、「二つ」のまま「規則」を提示する方が（自身にとって）得策だと、判断

するかもしれない。つまり「あなた」は、規範的な理論のみでなく、(応答が為される実際の状況での)「応答者」の心理的な側面をも、考慮に入れなければならないようになってくるのである。

これに続く段落は二文のみからなり、この小節の末尾であり、これを次に引く。

The appropriate incentive for you to offer a respondent for his opinion depends not only on the importance for you of obtaining that opinion with a specified accuracy but also on the difficulty for the respondent in obtaining it from himself. This makes the choice of a scoring rule designed to evoke the right degree of effort from the respondent on the various components of his task particularly subtle.

「あなた」が「応答者」に対して、彼から彼自身のオピニオンを抽出する際に、いかなる適切な刺激を与えれば良いのかは、ある特定の精度を伴って問題の(彼の)オピニオンを獲得することの(「あなた」にとっての)重要性のみでなく、彼が(彼自身から)問題のオピニオンを抽出することの困難さの程度にも依存しているのである。この彼自身が抱える困難さの問題は、「あなた」が彼からなんとか引き出そうとしている(オピニオン抽出のための)「努力」を、実際になんとか引き出すように(「あなた」が)造り上げたはずの、「得点規則」の性格をかなり微妙なものとする。彼は自身が立脚しているはずの多様な要素に基づいて、自身のオピニオンを抽出するのだが、「あなた」が与えている(「努力」のための)刺激的な誘導にもかかわらず、抽出作業そのものの困難さ故に、言わば「内在的な困難」の故に、(彼は)「あなた」の需要を満たすことができないかもしれない。

12. 割安の「得点規則」などない

第11小節は There are no bargain scoring rules. である。「得点規則」とは、応答者に対する「支払いの計画」である。そこで、応答者からの(少なくとも質問者にとっては)適切な情報を、できる限り低い「支払い」で獲得しようと、質問者が試みることは、異様ではない。だが、「得点規則」の性格からして、そのような割安の「規則」があるとは思われないというのが、サヴェジ氏の立場である。

この小節の二番目の段落を引くと次である。

The respondent will presumably work for what he regards as an acceptable wage. In the presence of a scoring rule, he will perceive his wage as random with an expectation that can be adjusted by adding a constant to the scoring rule. So, within the linearity approximation, the respondent does not charge extra for submitting to the scoring rule, and it therefore seems roughly reasonable to reckon that the questioner is not charged for it.

「応答者」は、受容できると思える「賃金」が得られると見込めば、なんとか自分の仕事をこなしてくれるであろう。但し、ここでの「賃金」とは、実際にはランダムなものであり、しかも、ある一つの(「応答者」にとっての)期待値を伴っている。この「期待値」は、「得点規則」に一定の金額を(「質問者」が)追加することによって調整できるのであり、「質問者」は、「応答者」の欲求を考慮して、「得点規則」を調整する。このような(調整された)「得点規則」に(「応答者」が)従うことが、調整を行う「質問者」にさらなる負担をもたらすわけではない。但し、調整の範囲が、「貨幣の効用」が線形的であると見なせる程度にとどまるとしておく必要がある。

これに続く段落を引く。

The higher the amplitude of the scoring rule, the more incentive it gives the respondent to reply with care. On this account, the respondent might in principle come to insist on a higher mean wage for facing a highly fluctuating scoring rule, and this could tend to deter the questioner from using high amplitudes. But the important practical limitation on amplitude would seem to be the need to avoid the distortion of response induced by the nonlinearity of the respondent's utility.

「得点規則」は「ランダムな支払い」を提示しているので、そこには「支払われる」金額の幅がある。この「金額の幅」をより大きくすることによって、「質問者」は「応答者」の「得点規則」に対する努力及び注意を、集中させ強めることが多分できるのである。だが、この「幅」が拡大するに従って、「応答者」は自身の「損失」を回避しようとして、より高い「平均的な賃金」を欲することとなり、結局、「質問者」側の負担が増大することと多分なるのである。但し、「幅」の増大に伴う実際上の難点は、「問題の「金額の幅」の範囲では（「応答者」の）「効用」は、近似的に線形である」という、「適正な得点規則」の（その成立のための）前提が揺らぎ、「（応答者）の」「貨幣の効用」の非線形性が、現実の実験の場にまで影響をもたらすことなのである。なお、この段落でこの小節は終わっている。

13. ビジネスの分担

第12小節は Business sharing である。冒頭の段落を引く。

In common sense, we feel without any overt reference to economic models that some responses are not so wrong as others and ought not to be so heavily penalized. If

a respondent is pretty sure that the home team will win, and there is in fact a tie, then he is perhaps not so wrong as if it had lost. Scoring rules reflecting this idea have been sought and easily found. See, for example, [21,53]. One interesting way to adjust the rewards and penalties of the respondent to the interests of the interrogator, which was brought out in a dramatic and radical form in McCarthy's [37] pioneering note on proper scoring rules, is to give the respondent a fractional interest in the business involved. To illustrate with an overidealized example, an oil prospecting company could give its geologist a small fraction of all profits and losses with the understanding that all decisions in the business would be made using the geologist's personal probabilities about geological uncertainties.

経済的な状況に明白に言及せずとも、さらには経済的な解釈ができるモデルを露骨に持ち出さずとも、「応答者」が示す諸反応の間に「自然な」格差があるとした方が、「まとも」である場合がある。例えば、「応答者」が、「こちらのチームの方が紛れもなく勝つだろう」と判断して、しかし実際には「引き分け」であった場合、仮に「負け」であった場合と比較して、「引き分け」の方が（「応答者」が荷う）「罰」は、相対的に「少ない」とすることは、我々の「常識的な感覚」に合致しているであろう。なおサヴェジ氏は、このような「常識的な感覚」に合致する「得点規則」が提案されていることを注意している。しかしさらに彼は、「応答者」の荷う「報酬や罰」を、「質問者」の利益との関りで調整して行くやり方があることを、述べている。つまり、問題となっているビジネスがもたらすであろう「利益」の一部を、一定の手順に従って「応答者」に分け与えることを、「得

点規則」に組み込むのである。例えば、「質問者」及び「応答者」として各各「ある石油探査会社」及び「ある地質学者」を想定する。この場合、「問題となっているビジネスの地質学的な不確定性に関しては、「その地質学者」の個人的確率を全面的に利用する」と両者が了解した上で、「その石油探査会社」は、問題のビジネスがもたらす全体的な「利益及び損失」の(相対的には少ない)一部分を、「その地質学者」に(一定の手順に従って)分け与えることとするのである。このような取り決めを了解した上で、「その応答者」は、自身から「確率」を抽出し、その「確率」を正しく「質問者」へと伝えるわけである。

この段落の後に括弧つきの追加の段落があるので、これを引く。

(It would be interesting to consider with some care the respects in which such an example is realistic and unrealistic, but I can only go a step or two in that direction here. Stock in the company would seem to give the geologist an interest in reporting his probabilities honestly if he could be assured that they would be adopted as the personal probabilities of the management for the events concerned. But this incentive may not be fully in harmony with the expert's incentives to appear worthy, as opposed to simply being worthy, of retention and promotion—a complication that affects not only business sharing but all applications of proper scoring rules to a professional expert. Business sharing does not present the expert with an explicit scoring rule in any business complicated enough to provide a more than mechanical role for managers, in particular in any business in which there are other uncertainties than those about which the expert is consulted,

but an implicit rule is as effective in principle as an explicit one.)

上で述べたように(「応答者」に)ビジネスの「利益」を分担させることによって、言わば潜在的に「得点規則」を(「質問者」が「応答者」に)あてはめるやり方がどこまで現実的なかは、個々の事例によって変動するであろう。「地質学者」が提示する「確率」を、問題となっている諸事象に対して、常に当てはめる」ことを前提とすれば、「その地質学者」が「その石油探査会社」の株を分有することは、彼が誠実に自身の「確率」を告げることを(彼へともたらされるであろう「利益」によって)、誘導することとなるであろう。つまり、自身の「確率」の使われ方を承知している株の所有者である彼は、誠実に述べるのが、自身の期待「利益」を最大化することを了解するであろう。だが通常は、企業の経営者は、エキスパートが問題とする諸事象よりも、さらに多様な事象を問題としており、エキスパートの言い分を処理する機械ではない。一方、知的なエキスパートは、単なる「利益」のみに反応するのではなく、(企業における)自身の地位なども問題とするであろうから、彼も「利益」のみに反応する機械ではない。しかし、「潜在的な「得点規則」」が、「明白な様式で提示される「得点規則」」と同様に有効に機能することが、十分に期待され得るのである。

これに続く段落を引く。

Long ago, Gauss [25, Sec.6] proposed that economic losses (such as those in a game of chance) provide a good model for the incentive to estimate accurately even in the most academic contexts. Decisively to uphold or to overthrow this suggestion does not seem possible. Personally, it appeals to me. Correspondingly, when we say that a tie should not be regarded as so distant

from a win as a loss would be, I am inclined to think that that is because we have in mind various uses for sport forecasts in which the penalty is less for one kind of error than another. Of course, the penalty need not be a monetary one; it might involve, for example, loss of social prestige. Fortunately, the elusive question of whether all that is good and bad about a forecast can ultimately be referred to profit and loss in economic decisions, sufficiently, widely interpreted, need not be resolved in order to show the interest and utility of viewing proper scoring rules as a share in a real or a fictional business. For the business model is certainly a mathematically general model for all proper scoring rules and a fertile point of view for the generation of proper scoring rules that penalize some errors more than others.

未知かつ固定されている「真の」値を推定する作業においては、推定のために導入される値（つまり「推定値」）と「真の」値との、何らかの「隔たり」が問題となるのだが、この「隔たり」を、何らかの「経済的な損失」と結びつけて考察する流儀が存在し得る。ここで、「損失」は、問題の推定を行うその本人が荷うこととすると、その本人は、自身の期待される「損失」をできる限り少なくしようと試みるであろうから、推定の問題を「経済的な損失」と結びつけて考察する流儀は、推定を行う者に対して、「より正確な」作業を行うように刺激を与えていることとなる。ここで注意すべきなのは、推定の問題が（通常の意味でのビジネスとはことなる）全く学問的なものでも、「経済的な」モデルによるこの接近が可能なのだということである。このような可能性は既に Gauss が指摘しているのだが、サヴェジ氏はこのような流儀を重視している。彼が注意しているよう

に、この流儀を（合理的な根拠によって）棄て去ることなどできないであろう。一方我我は、しばしばスポーツとの関りで、明白に意識しているか否かは別として、予測の「誤り」に序列をつけている。つまり例えば、「こちらが勝つ」と予測して、実際には「引き分け」であった場合と、実際には「負け」であった場合とでは、前者の「誤り」がもたらす「ペナルティー」は後者よりも「少ない」と判断すること、などである。なお、勝負の状況が「得点」でより細かく表現されれば、より細かく「ペナルティー」が設定できるであろう。つまり我我は、明白に意識はせずとも、「経済的」と形容してもそれほど不自然ではない「ペナルティー」を使っているのである。また、「適正な得点規則」を、現実の（あるいは架空の）ビジネスでの（「利益」の）分担を定めるものと、見なし得るであろうし、一方、ビジネス・モデルは、「決定」を下す者が犯す「誤り」に序列をつけることによって「決定者」に「ペナルティー」を課す「規則」を提示することによって、「適正な得点規則」を潜在的に定める数学的で一般的な「得点規則」と見なし得るであろう。

さらにこれに続く段落を引く。

The technical point that every proper scoring rule can be viewed as a share in a business and that every such share leads to an at least weakly proper scoring rule should be appreciated. Section 7 makes these points clear. Every strictly proper scoring rule amounts to the possibility of choosing among acts, only one of which is appropriate to each system of personal probabilities. Conversely, a person knowing that an act in a specified economic situation is to be chosen for him in accordance with his announced system of personal probabilities will have no incentive to announce a false one. However, if the convex function arising

from the family of acts has flat places—technically, is not strictly convex—then the person has no positive incentive to distinguish among certain systems of probabilities, and the scoring rule is only weakly proper.

ここでサヴェジ氏は、Elicitation (1971) の第7節の内容に言及している。つまり、「適正な得点規則」の各々は、あるビジネスでの（「利益」の）ある分担と見なし得るし、そのような「分担」の各々は、「弱く適正な得点規則」へと通じている、というのである。「行為」の間での選択が問題となっており、「決定者」の可能な個人的確率の各各に対応して、唯一の「行為」が、「決定者」の期待損失を最小化するという状況は、「（狭義に）適正な得点規則」がもたらすものに他ならない。一方、特定の経済的状况を想定した上でだが、自身が告げる個人的確率に従って、その告げられた「確率」によれば（その当人にとって）最も期待損失が少ない「行為」が選ばれるとすれば、このことを知っているその当人は、「偽りの確率」をわざわざ述べることへの積極的な動機づけを、持たないことであろう。だが、可能な「行為」の全体がもたらす凸関数が「狭義に凸」ではなく、「たいらな」部分を持つならば、「決定」を下す当人は、ある複数の可能な「確率」をわざわざ区別するための積極的な動機づけを持たないであろうから、このような状況がもたらす「得点規則」は「弱く適正」である。

さらに（一文のみから成る）これに続く段落を引く。

Insofar as business sharing is a practical method of elicitation, it makes possible the use of a proper scoring rule even in those situations stressed at the end of Subsection 10.3 in which arbitrary proper scoring rules cannot be implemented because some of

the conditional probabilities to be elicited will not be tested, depending on the opinions expressed by the expert.

第3小節の末尾では、（エキスパートから）抽出される（条件つき）「確率」の幾つかは、エキスパートの提示するオピニオンを考慮して実際の場合で試されることがないであろうから、「いかなる「適正な得点規則」でも実行可能だ」とはならないであろう、との趣旨の発言をした。だが、その問題の場合でも、ビジネスの分担が実際的な（「確率」の）抽出方法をもたらすのならば、「適正な得点規則」が利用可能であると見て良いであろう。

さらに（この小節の末尾である）これに続く段落を引く。

The parallelism between implementing a scoring rule by business sharing and the rough rule for rating experts according to how their advice would have affected your business in the past (discussed in Subsection 10.8) is evident, but the two things must not be confused. In particular, the rating rule can be used regardless of what scoring rule if any is used to elicit the opinions.

第8小節では、エキスパートたちの助言が、（仮にそれらが利用されたとして）どの程度まで「あなた」のビジネスに「利益」をもたらしたであろうか、ということに着眼して、彼らを格づけするやり方を議論した。このやり方と、ビジネスの分担によって「得点規則」を遂行する流儀とは、見かけ上は似ていても本質的に異なるのであり、実際、「格づけ」のやり方は、いかなる「得点規則」が（仮にそのような「規則」があるとして）採用されたのかには依存せずに、利用し得るのである。

14. 安楽椅子心理学

第13小節はSome armchair psychologyである。冒頭の段落を引く。

Consider now a subject quite untrained in personal probability and the theory of scoring rules whose only incentives are provided by the scoring itself. It is questionable whether, in any serious application, this ought to be allowed to happen. Though any strictly proper scoring rule is a sufficient guide for an ideally intelligent Robinson Crusoe no matter how uninstructed, can we expect real people to respond well with no other coaching than is provided by the scoring rule itself or even by extended experience with the scoring rule? And even if investigation should yield a somewhat affirmative reply, is there any point in withholding instruction? Whatever the answers, it does seem stimulating to speculate on what kinds of scoring rules presented in what way would most nearly operate on naive subjects as all strictly proper scoring rules are supposed to operate on sophisticated ones, and this should lead to ideas of practical value for subjects of intermediate sophistication.

理想的に知的な「応答者」は、彼が個人的確率や「得点規則の理論」について何の訓練も受けていなくとも、「得点規則からの得点」と「得点規則についての経験」とで、自身の人生を「満足して」やっけて行けるはずである。このような主張は多分現実的ではあるまい。しかし、わずかでも「現実味」があるとしても、「応答者」に（個人的確率に基づく）「得点規則の理論」を学ぶことを控えさせる理由にはならないであろう。だが、「熟練した応答者」に対しては、全ての「(狭義に) 適正な得点規則」がうまく

作用する」とした上で、何の訓練も受けていない「応答者」に対して、いかなる「得点規則」をいかなるやり方で提示すれば効果的なのかを、考察することは、興味あることであるし、このような考察は、訓練の程度が中間的な「応答者」にはどのように対処するのかについて、実際的な知恵をもたらすかもしれない。

これに続く段落を引く。

In the first place, the subject must understand the scoring rule. If it makes explicit reference to logarithms, or even to squares, most ordinary people will not understand it at all; and even those with mathematical training may not be nearly apt enough at calculation to use the rule effectively. This is an important reason to present the rule through some vivid tabular or graphic device, which could, for example, take the elaborate form of conversational-mode digital computation, or more simply of some slide-rule device such as those of the Shuford-Massengill Corporation (Lexington, Mass.), or perhaps tabulation of the scoring rule itself, possibly very boldly rounded [14].

とにかく被験者、つまり「応答者」は、提示されている「得点規則」とはいかなるものなのかを、何よりもまず理解しなければならない。だがここで早速問題に突き当たる。つまり、「得点規則」が対数や、それどころか自乗に、明白な様式で言及しているのならば、通常の人人は、その「規則」を理解してくれないであろう。それどころか、通常の数学的な訓練を受けたはずの人人でさえ、そのような「規則」を、定量的にしっかりと理解できるかどうかは疑わしい。そこで、「得点規則」を、直接的ではなく、表やグラフを使って提示することが試みられることとなる。なおサヴェジ氏は、対話式のデジタル計算機や、Shuford-Massengill社の

計算尺などに言及している。しかし彼は、「得点規則」そのものを大まかな表で提示するという、de Finetti の流儀の有用性に気づいている。

さらに（この小節の末尾である）次の段落を引く。

Perhaps it is helpful to a subject responding about a partition of possible events if the economic consequence of his response is a function only of the element of the partition that happens to obtain. This condition imposes no constraint at all for two-fold partitions, but for n -fold partitions with $n > 2$, it leads to the logarithmic scoring rules of Subsection 9.4. The possible advantages of the simplicity might often be outweighed by the inappropriateness of a symmetric scoring rule in asymmetric situations or of a scoring rule that lays emphasis on the correct elicitation of small probabilities. I have sometimes heard the possibility that a subject responding to a logarithmic scoring rule could be subjected to an infinite (or at any rate, unlimited) penalty raised as an overwhelming objection. This possibility does of course imply that the method cannot be applied literally, but approximate applications, in which the subject is not allowed to name probabilities less than, say, 10^{-3} suggest themselves. And, as mentioned in Subsection 10.5, obtaining probabilities very close to 0 by means of direct incentives does not seem practical by any scoring rule.

「世界」の分割を事象の（「長さ」が2以上の）系列として捉え、どの事象が実際に「通用する」のかは、少なくとも「応答者」にとっては「ふたしか」であるとし、さらに、「応答者」に（結果として）もたらされる経済的結果は、

「実際に「通用する」こととなる事象に対応する（「応答者」が定める）「値」のみに依存する」とすると、「長さ」が3以上であれば、第9節の第4小節で示したように、「応答者」が従う「得点規則」は、対数的なものとなる。だが、「得点規則」の単純さがもたらす利点よりも、非対称な状況に「対称的な得点規則」をあてはめたり、小さな「確率」を正確に抽出することを強調する「得点規則」を採用したりすることのほうが、「確率」の抽出実験にとって（しばしば望ましくない結果をもたらすので）重大であるかもしれない。なおサヴェジ氏は、「対数的な得点規則」への強い異議があることを承知している。つまり、「対数的な得点規則」を提示される「応答者」は、「制限のない罰」にさらされる可能性と直面するであろうというのである。このような論難に対して、彼は、数学的に導出される「得点規則」は、現場では、ある種の近似として取り扱うべきだと判断している。なお、第5小節で言及したように、微小な確率を直接的に「得点規則」を用いて抽出するやり方には、実際上はかなりの無理がある。

15. 学校と「適正な得点規則」

第14小節は Proper scoring rules in school で、冒頭の段落は次の二文である。

Proper scoring rules hold forth promise as more sophisticated ways of administering multiple-choice tests in certain educational situations [14,49]. The student is invited not merely to choose one item (or possibly none) but to show in some way how his opinion is distributed over the items, subject to a proper scoring rule or a rough facsimile thereof.

「適正な得点規則」とは、「「応答者」が、（その「得点規則」に従う場合の）「得点」に対応する（自身の）効用の（自身の）「期待値」を、

(あるいは「期待損失」を), 最大化 (あるいは最小化) することを欲する場合に, 彼は (自身の) 本来の「確率」を告げざるを得ない」という状況に, (「回答者」を) 置く「規則」である。このような「規則」を, 学校での試験に応用しようと言うのである。つまり, 選択式の問題で, 「回答者」が (自身が正しいと思う) 唯一の選択肢を選ぶ, さらに「何も選ばない」ことをも「選択肢」の一つとする, 通常の状態において, 「唯一」の「答え」を選ぶのではなく, 中間的な「値」を各選択肢に配分するように, 「回答者」に要求するのである。この際, 「値」の分布に対応して, 「適正な得点規則」による「得点」が「回答者」にもたらされるのだが, その点数配分のありさまを, おおまかに (但し, わかりやすく) 「回答者」にあらかじめ提示する必要があるであろう。ここで, 「回答者」が配分する「値」が, 「その選択肢」が正しいことの (彼にとっての) 「確率」の近似値である。

これに続く二文から成る段落を引く。

Though requiring more student time per item, these methods should result in more discrimination per item than ordinary multiple-choice tests, with a possible net gain. Also they seem to open a wealth of opportunities for the educational experimenter.

「適正な得点規則」を, (「選択肢」としての) 「項目」を選択する問題に当てはめると, 各「項目」に対する (「学生」が必要とする) 時間は増加するであろうが, より精密に「項目」どうしを (「学生」の立場から) 識別できるようになってくるであろう。この場合, 従来のただ一つだけを選択するやり方よりも, 「学生」自身の期待効用は, 通常は増加するのである。

これに続く段落を引く。

Above all, the educational advantage of training people—possibly beginning in early childhood—to assay the strengths of their own opinions and to meet risk with judgment seems inestimable. The usual tests and the language habits of our culture tend to promote confusion between certainty and belief. They encourage both the vice of acting and speaking as though we were certain when we are only fairly sure and that of acting and speaking as though the opinions we have were worthless when they are not very strong.

ここでサヴェジ氏は, 個人的確率という事柄を学ぶことの重要性を主張している。「個人的確率の概念」そのものではなく, 個人的確率とはいかなる事柄なのかを, 日常の体験を通して「学ぶ」ことの重要性を強調しているのである。自身のオピニオンの強さを測定し, リスクに対して「自身の判断」で対処してゆくことを, 幼年期の初期から学ばせることは, その子が生きて行く上で極めて有益なことであろう。彼は, 教育の過程に個人的確率をぜひとも組み込むべきだと判断し, しかもさらに, 「確率」教育の社会への (健全な) 影響を力説する。(恐るべきことだが) 普通に行われている「試験」や言語的習慣は, 確定性と (個人の) 信念とを混同させることを, 促進させているのである。たとえば, 「かなり確かだ」と感じているだけであるのに, それが「確実である」かのように振舞い, そして話すのである。しかも, さらにまずいことには, オピニオンが「非常に強い」ものでないとすると, そのオピニオンは価値がないと断定するかのように, 振舞い, そして話すのである。このような社会的な悪習を減少させる方向に個人的確率は機能すると, サヴェジ氏は「希望」している。

これに続く段落を引くが, これが末尾の段落である。

Effects of nonlinearity in educational testing deserve some thought, but presumably nonlinearity is not a severe threat when a test consists of a large number of items. One source of nonlinearity that has been pointed out to me is this. A student competing with others for a single prize is motivated to respond so as to maximize the probability that his score will be the highest of all. This need not be consistent with maximizing his expected score, and presumably situations could be devised in which the difference would be important.

「得点規則」を教育の現場での「試験」に導入する際に、その「規則」の「得点」に対する（「応答者」としての「受験者」の）「効用」が問題となる。「効用」が、問題となる「得点」の範囲において、線形的な状況から大きくずれれば、「受験者」は本来の（自身の）「確率」ではない「値」を、（「選択肢」としての）「項目」に配分することとなり、「線形的な状況」を前提とした第9節までの議論は当てはまらなくなる。しかし、「項目」が多数設定されている状況では、この「線形的な状況」を想定しても良いであろう、というのがサヴェジ氏の見込みである。だが、「試験」が、ただ一つの賞を目指して行われる競争試験である場合、「自身の得点が全員の内で最高である確率」を、最大化するように、「値」を配分する「受験者」がいるとしても不思議ではない。このような「最高得点の確率」の最大化と、自身の期待「得点」の最大化とが、同値である保証はない。なおサヴェジ氏は、これら二つの最大化が無視し得ない程に異なる場合があるであろうと、推測している。

16. おわりに

サヴェジ氏の Elicitation (1971) の第10節を通読してきたわけだが、sidelightsではすま

されない内容であることがわかる。また、彼自身の態度にも十分に注目すべきである。Elicitation (1971) の冒頭から第9節までは、個人的確率の抽出に関する数学的な議論であり、これらの議論で論述を終了したとしても、元来数学者であるサヴェジ氏が非難される筋合いはないであろう。だが彼はさらに一節を追加せざるをえなかった。何故か。結局彼は、自身の「思索の様式」に忠実であったということであろう。この「思索の様式」とは、彼自身の「数学」と呼んでも良いものである。彼は自身の数学的議論が、「現場」とどのように関るのかを、真剣に問わざるを得なかったのである。意思決定の「現場」に「自身」が関る場合、さらには「自身」が正に当事者である場合、自身が展開した「数学」は、どのような意義を持ち得るのであるかと、彼は真摯に問うのである。つまり、通常の「数学」が「証明終」となっても、彼の「数学」は「終わらない」のである。彼は、自身の言い分にとって（少なくとも一見すると）都合が悪いような状況を、敢えて探查し、自身の数学的議論を執拗に「ためす」のである。このような「ためし」の場を、彼自身の「数学」は持っているのである。

数学的議論は「知」の場で展開され、一方「現場」はあくまでも「行」の世界である。サヴェジ氏の「思索の様式」は、「知」の産物を「行」の世界で執拗に「ためす」ことによって成立しており、彼の「数学」は、このような「ためし」の場を持っている。

2006年3月13日(月)

参考文献

Raiffa, Howard, *Decision Analysis: Introductory Lectures on Choices under Uncertainty*, Addison-Wesley, Reading, MA, 1968.

Savage, Leonard Jimmie, *The Foundations of Statistics*, Wiley, New York, 1954. *Second Revised Edition*, Dover, New York, 1972.これは「基礎論」であり、統計学へのサヴェジ氏の偉大な貢献である。なお、園(2000年6月)(あるいは園(2001年12月)の第2章)にサヴェジ氏の略伝がある。

Savage, Leonard Jimmie, "Elicitation of personal probabilities and expectations," *Journal of the American Statistical Association*, 66, 783-801, December 1971.「個人的確率の抽出」に関する真剣な考察であり、「確率」に関する古典的傑作である。「論文集」Savage(1981)に収録されている。

Savage, Leonard Jimmie, *The Writings of Leonard Jimmie Savage—A Memorial Selection*, prepared by a Committee (W. H. DuMouchel, W. A. Ericson (chair), B. Margolin, R. A. Olshen, H. V. Roberts, I. R. Savage and A. Zellner) for the American Statistical Association and the Institute of Mathematical Statistics, Washington, D. C., 1981. サヴェジ氏の論文集である。

園 信太郎,「サヴェジ氏の略伝」,『経済学研究』(北海道大学),第50巻第1号,164(164)–180(180),2000年6月。これはサヴェジ氏の論文集(1981)に基づく「略伝」だが、彼の人柄を知る助けになるかもしれない。これは園(2001年12月)の第2章に収められている。

園 信太郎,『サヴェジ基礎論覚書』,岩波出版サービスセンター,東京,2001年12月20日。「基礎論」への要約,注釈,及び「読み」を提示している。