Cognitive Studies, **18**(1), 173-189. (March 2011)

●研究論文●

基準率無視と自然頻度の幻想: 等確率性仮説に基づく実験的検討

西田 豊・服部 雅史

We classify the explanation of the base rate fallacy into three positions: (1) the neglect view, which argues that the neglect of base rates causes the fallacy; (2) the frequency view, which claims that the fallacy will disappear if information is presented in the natural frequency format instead of the probability format; and (3) a family of theories that focus on mental representations of the task structure. In this paper, three experiments examined the validity of the equiprobability hypothesis, which is one of the theories that emphasize mental representations of the task structure. In Experiments 1 and 2, although we used tasks in which the neglect view predicted high performances, correct answers remained infrequent. In Experiment 3, facilitation by the natural frequency format could barely be distinguished from the effect predicted by the equiprobability hypothesis. Consequently, we suggested that the neglect view was inappropriate and the frequency view could be reinterpreted according to the equiprobability hypothesis.

Keywords: the equiprobability hypothesis（等確率性仮説）, Bayesian probability inference（ベイズ的確率推論）, the base rate fallacy（基準率錯誤）, natural frequency（自然頻度）, probability judgment（確率判断）

1. はじめに

不確実な状況下において人間がどのように判断を行っているかについて，これまで多くの議論がなされてきた．Tversky & Kahneman (1974) は，人の推論様式が規範的でないことを強調した．たとえば，次のような確率推定課題 (Eddy, 1982; Gigerenzer & Hoffrage, 1995) において，人の推論結果は正解から大きく逸脱することが知られている (レビューとして Barbey & Sloman, 2007; Koehler, 1996).

> 乳がんの検査に参加した 40 代の女性が乳がんである確率は 1% です．もし女性が乳がんなら，その女性が検査で陽性になる確率は 80% です．女性が乳がんでなくとも，検査で陽性になる確率は 9.6% です．検査に参加した 40 代のある女性が陽性でした．この女性が本当に乳がんである確率はどのくらいでしょうか？
> ＿＿＿%.
>
> (正解：約 7.8 %)

乳がんの検査に参加した 40 代の女性が乳がんであるという仮説を H，検査で陽性であるというデータを D とすると，課題文中に与えられているのは，何もデータがないときにこの女性が乳がんである確率（基準率）$P(H) = .01$，乳がんのときに陽性というデータが得られる確率（尤度）$P(D|H) = .80$，乳がんでないときに誤って陽性というデータが得られる確率 $P(D|\bar{H}) = .096$ である．この課題は，次のベイズの定理を使って解くことができる．

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)} \quad (1)$$

この式を適用するためには $P(D)$ が必要になるが，$P(D) = P(D|H)P(H) + P(D|\bar{H})P(\bar{H})$ ゆえ，$.80 \times .01 + .096 \times .99 = .103$ と算出できる．

Illusions of Base Rate Neglect and Natural Frequency: An Empirical Examination Based on the Equiprobability Hypothesis, by Yutaka Nishida (Osaka University), and Masasi Hattori (Ritsumeikan University).

この課題の正解は約 7.8% であるが，回答者の多くは 80% とベイズ解からかけ離れた値を答えてしまう．このような課題（以下では基準率課題と呼ぶ）において広く観察される正解と回答とのズレは，基準率錯誤 (base rate fallacy) と呼ばれている (Bar-Hillel, 1980)．基準率錯誤の説明は，大きく 3 つに分類可能と考えられる．1 つ目は，基準率無視 (Tversky & Kahneman, 1980) の考え方（以下では，基準率無視説，または単に無視説と呼ぶ）で，人は基準率を無視するために錯誤が起こるとするものである．2 つ目は，自然頻度仮説 (Gigerenzer & Hoffrage, 1995) である．この考え方によれば，情報の提示形式が確率であるから間違えるのであって，自然頻度 (natural frequency) の形式で提示すれば正答できる．3 つ目は，課題構造の心的表象に着目する考え方 (Evans, Handley, Perham, Over & Thompson, 2000; Hattori & Nishida, 2009; Sloman, Over, Slovak, & Stibel, 2003) である．課題の正答には，課題構造とその構造が心的にどのように表象されるかが重要であるとする．ここでの表象とは事象に関する主観確率のみではなく，事象間の関係を含めたものを意味する．以下でこれらの説を概観する．

1.1 基準率無視の神話と自然頻度仮説

まず，基準率無視説において，「無視」というのがどのような意味で用いられているのか明確にしておきたい．Tversky & Kahneman (1980) は，基準率錯誤について，次のオッズ形式のベイズの定理を用いて説明した．

$$\frac{P(H|D)}{P(\bar{H}|D)} = \frac{P(D|H)}{P(D|\bar{H})} \cdot \frac{P(H)}{P(\bar{H})} \qquad (2)$$

右辺は，陽性というデータが得られる尤度のオッズと基準率のオッズの積になっている．尤度のオッズは検査の信頼性を示し，この値は大きい (.80/.096 = 8.3) が，基準率オッズは非常に小さい (.01/.99 = .01) ため，結果として，仮説の事後確率のオッズ（左辺）は小さくなる．このことを Tversky & Kahneman (1980) は，基準率がデータの信頼性より「極端 (extreme)」であるために，仮説の確信度がデータと逆になると表現した．すなわち，データは仮説を支持する方に傾いているが，基準率がそれ以上に仮説を打ち消す方向に傾いているため，結果として仮説は不支持の方向に傾くことになる．ところが，「圧倒的多数の参加者は基準率を考慮することに完全に失敗する」[1] (p. 63) ために，仮説が支持される（事後確率が高い）と考えてしまう．

つまり，基準率無視というのは「基準率の極端さが無視される」という意味であり，基準率の極端さとは，基準率 $P(H)$ が非常に小さいこと（よって，その否定 $P(\bar{H})$ の確率が大きいこと）を指す．したがって，基準率の無視とは「$P(H)/P(\bar{H}) \simeq 1$ と仮定すること」と考えられる．$P(H)/P(\bar{H}) = 1$ であれば，式 (2) において基準率オッズの項が存在しない（無視する）ことと同じ意味になることによると考えられる．

Tversky & Kahneman (1980) も基準率が常に無視されるわけではないことは認めていたが，Koehler (1996) は，それまでの基準率課題の実験結果を網羅的にレビューして，「基準率を完全に無視する人がいるという結果を示した研究はほとんどない」[2] と述べ，基準率無視を「神話 (myth)」と称した (p. 3)．これは，基準率無視説が間違っているか，あるいは少なくとも不十分であることを端的に示している．ところが，これまで，基準率課題を材料にした研究は数多く実施されてきたものの，極端さが無視されるという意味において本当に基準率が無視されているのかどうかを直接的に検討した実験は見当たらない．したがって，この点について実験的に検討する価値がある．

一方，基準率課題に正答できないのは，課題文で提示される統計情報が確率形式であるからという主張もある (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995)．Gigerenzer & Hoffrage (1995) は，人は自然頻度によって思考するように調節されているとして，百分率や相対頻度で情報を与えることとは，生態学的妥当性が低いと主張し，統計情報が自然頻度で表現された次のような課題を用いてその妥当性を検証した．

> 乳がんの検査に参加した 40 代の女性の 1000 人に 10 人は乳がんです．乳がんである女性の 10 人のうち 8 人は，検査で陽性になります．乳がんではない 990 人のうち 95 人も，検査で陽性になります．検査に参加した 40 代のある女

1) 原文では “The overwhelming majority of subjects fail altogether to take the base-rate into accout.”
2) 原文では “Subsequent research failed to support a complete neglect of base rates.”

性が陽性でした．この女性が本当に乳がんである確率はどのくらいでしょうか？ ＿＿＿人中＿＿＿人．

(正解：103 人中 8 人)

確率表現を使ったときには 16% だった正答率が，自然頻度表現を用いると 46% にまで上昇したと報告した（実験 1, p. 693）．彼らは，自然頻度表現を用いると事後確率の計算が簡単になるため正答率が上昇すると説明した．しかし，以下で述べるように，自然頻度仮説を疑問視する研究も多くある．

1.2 課題構造と等確率性仮説

Evans et al. (2000) は，自然頻度表現は必ずしも正答率を上昇させるわけではないとした．正解率が上昇するのは頻度表現そのものの効果ではなく，要素間の関係がよりわかり易くなるためだと主張した．また，Sloman et al. (2003) は，集合関係が入れ子構造になっていない課題はより難しいと予測した．これらは，課題で与えられる集合・要素の構造に着目し，それらがどのように表象されるかが重要であることを説いた研究である．

Hattori & Nishida (2009) は，基準率課題における直感と正解の乖離を等確率仮説によって説明した．等確率性仮説 (equiprobability hypothesis) とは，人は焦点となる 2 つの事象 (H と D) の等確率性 $P(H) \simeq P(D)$ を暗黙のうちに想定するとするもので，条件文推論や因果帰納，仮説検証の文脈で検証されている (Hattori, 2002; Hattori & Oaksford, 2007; 田村・服部・三輪, 2010)．この説も，課題構造と表象構造に焦点をおいた研究と位置づけることができる．

Gigerenzer & Hoffrage (1995) らは，基準率課題が難しい 1 つの要因として，事後確率を計算する際に必要となる $P(D)$ が課題文中に直接与えられないことを指摘した．確かに，数学の問題を解くときのように，紙と鉛筆を使ってベイズ計算を顕在的に実行する場合を考えれば，$P(D)$ は式 (1) の分母に相当し，計算途中で必ず導出が必要になることから，その数値を問題文中に与えれば正解が促進されることは容易に想像できる．しかし，この課題の重要な論点は，単に正確な数値を導くのが難しいという点ではなく，直感が正解から特定の方向に，しかも大きくずれるという点である．その意味では，基準率課題の本質的な難しさの原因は別のところにあると考えるのが妥当であろう．つまり，課題の心的表象がどのように形成されており，また，なぜそのように形成されているのかを追求することが重要である．

基準率課題に用いられてきた課題には，冒頭の乳がん課題のほかに，有名なタクシー課題 (Tversky & Kahneman, 1982) や感染者課題 (Casscells, Schoenberger, & Grayboys, 1978) がある．Hattori & Nishida (2009) によると，これらの課題は共通した課題構造を持っている．それは，焦点となる 2 つの事象の周辺確率 $P(H)$ と $P(D)$ の大きさが非常に大きく異なっている点である．図 1 に課題構造を示す．図 1-I では $P(H)$ が小さい (1%) のに対し，$P(D)$ は大きい値 (10.3%) をとっている．そのため，部分集合 $H \cap D$ が集合 H の中で占める割合は大きいのに対し，集合 D の中で占める割合は小さい．つまり，尤度 $P(D|H)$ は大きいが，事後確率 $P(H|D)$ は小さい．この課題構造が，基準率課題に正答することの難しさと深く関わっていると考えられる．このような特徴的構造，すなわち，$P(H) \ll P(D)$ という課題の確率構造を偏確率構造 (imbalanced probability structure) と呼ぶ．

偏確率構造を持った課題において，回答者が（誤って）図 1-II のように等確率性を仮定すると，尤度 $P(D|H)$ と事後確率 $P(H|D)$ が等しいことになる．すなわち，式 (1) の分母を等確率性仮定により $P(D) = P(H)$ とすると，$P(H|D) = P(D|H)$ となり，これは多くの回答と一致する結果となる．これまで，基準率を無視するために正しい値を答えることができないと考えられてきたが，実際は，等確率性を仮定してしまうためにあたかも基準率を無視したかのような値を答えてしまう，というのが Hattori & Nishida (2009) の論点である．

Hattori & Nishida (2009) は，2 つの実験によって等確率性仮説に基づく予測の正しさを確認した．実験 1 においては，実験参加者が想定する等確率性仮定を崩して課題の偏確率構造を正しく認識させることが，正答率を上げるかどうかが検討された．参加者の等確率性の仮定を崩すために，一般知識が利用された．病気（X 症候群）という原因事象に対する結果事象として，「陽性検査結果」の代わりに「咳」という代替原因の想定が容易な事象を用いるだけで，正答率が劇的に上昇した．また，偏確率構造を正しく認識している回答者ほど，課題に正答で

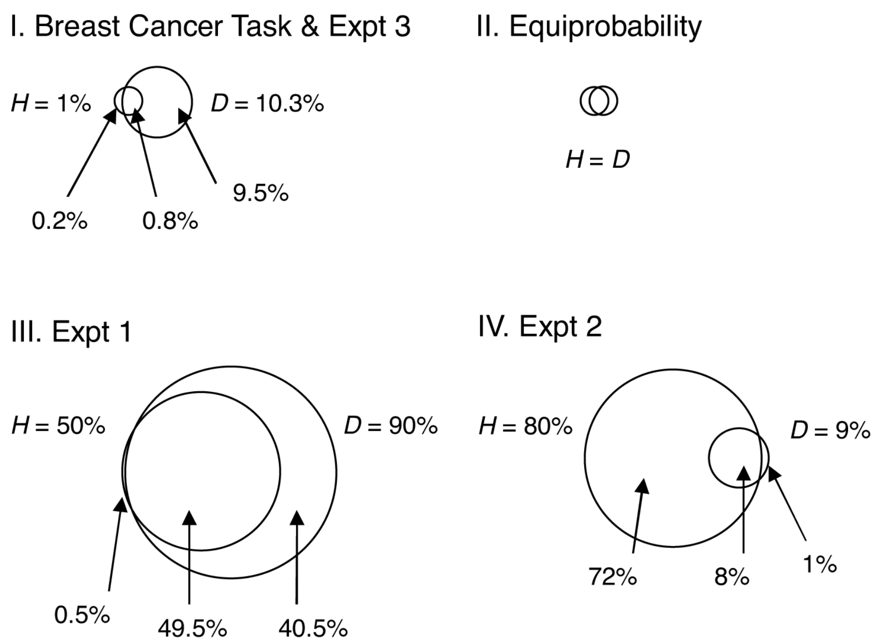


図 1 乳がん課題 (I) と実験 1 (III)，2 (IV), 3 (I) で用いた課題の確率構造，
および人々がデフォルトで想定すると考えられる等確率構造 (II).

きる傾向があることが明らかにされた．実験 2 においては，課題が困難なのは，暗黙になされる等確率性仮定と，明示的に与えられる課題構造が競合するためであることが，課題の偏確率の程度を操作することによって示された．

1.3 本論文の概要

本研究は，等確率性仮説 (Hattori & Nishida, 2009) の妥当性をさらに詳しく検証するために，3 つの実験を行った．他の説，特に，基準率無視説と自然頻度仮説では，実験結果の説明が困難または不十分であることを主張するものである．

実験 1, 2 は，基準率無視説の検討のために実施された．基準率課題に正答できない理由について，Tversky らは基準率が無視されるためだと考えたが，われわれは偏確率構造の認識が困難であるためだと考える．実験 1 では，基準率を無視することによって正解に導かれる課題を用いて，等確率性仮説の予測と比較しながら，無視説の妥当性を批判する．実験 2 では，基準率そのものの大きさ操作し，やはり無視説では正答率が上昇すると予測される課題を用いて，この説では実験結果の説明が難しいことを示す．

実験 3 は，自然頻度形式の情報がなぜ正答率を上昇させるのかについて検討した．自然頻度仮説では，人の思考様式に適応している自然頻度を使うために課題に正答できると主張するが， Hattori & Nishida (2009) においては，確率表現を使っても劇的な正答率の上昇が得られた．すなわち，人が基準率課題に正答できるか否かは情報の表現によらないことになる．それでは，自然頻度形式で情報提示すると正答率が上昇するのは，なぜであろうか．自然頻度形式と，等確率性仮説が主張する偏確率構造の理解は，どのような関係にあるのだろうか．この問題について，同一構造の課題を用いて，正答率と内的表象の違いを検討した．

2. 実験 1：基準率無視が正答に導くとき

基準率無視説によれば，基準率オッズの極端さ（値が 0 に近いこと）が無視されるために課題に正答できない．基準率オッズの極端さを無視するというのは，実質的には基準率オッズを 1 とみなすことに相当する．したがって，無視説が正しいとすれば，基準率オッズが 1 の課題，すなわち $P(H) = P(\bar{H})$ である課題を用いれば，基準率の無視が正解へ導くことになり，正答率が上昇するはずである．一方，

等確率性仮説によれば，基準率オッズが極端であるかどうかに関わらず，課題が偏確率構造を持ち，かつその認識が困難であるならば，正答率が下がると予測する．そこで，どちらの説が正しいかを実験的に検討した．

2.1　方　法

2.1.1　実験参加者

大学生 20 名が実験に参加した．実験は集団により行われた．1 名は $P(D)$ の回答に不備があったため，$P(D)$ に関わる分析では除外した．

2.1.2　BR 課題と Est-D/HD 課題

基準率オッズが 1 となるよう $P(H) = P(\bar{H}) = .5$ かつ偏確率構造を保つよう $P(D) = .9$ とし課題を作成した．実験 1 で用いた確率構造を図 1-III に示す．ストーリーは確率値が不自然にならないよう考慮し，以下を用いた．

> あなたは病院で働く薬剤師であると仮定してください．近年，X 症候群という難病に対する新薬が開発されました．あなたに求められているのは次の情報から医者が新薬を使ったかどうかを判断することです．
>
> X 症候群の患者に対して医者が新たに開発された薬を使う確率は 50% です．新薬を使ったならば，X 症候群の症状が改善する確率は 99% です．新薬を使わなかったとしても，X 症候群の症状が改善する確率は 81% です．ある X 症候群の患者の症状が改善しました．医者が新薬を使った確率は何 % でしょうか？　直感で答えてください．
>
> (正解： 55 %)

実験参加者には，基準率 (BR) 課題と 2 つの確率推定 (Est-D/HD) 課題が課せられた．BR 課題は，事後確率 $P(H|D)$ を推定するもので，実験参加者の最終出力（解）がベイズ解と一致しているかどうかを調べるための課題である．$P(H|D)$ を数値によって回答させた．Est-D/HD 課題は，周辺確率 $P(D)$ および同時確率 $P(H,D)$ をそれぞれ推定する確率推定課題である．課題確率構造の内的表象を探るための課題で，Hattori & Nishida (2009) で使用されたものと同じである[3)]（Est-D/HD 課題のサンプルを付録 B に示す）．

2.1.3　手続き

配布された冊子のフェイスシートには，課題文をよく読んでから答えること，確率を回答する際に計算の必要はなく直感で答えること，制限時間はなく自分のペースで答えること，3 つの課題すべてに答えることが教示してあり，実験開始前に実験者によって読み上げられた．回答時間は参加者ごとに異なるため，実験者に冊子を提出してもらい実験終了とした．

2.2　結果と考察

実験 1 で得られた主観事後確率 $P(H|D)$，主観周辺確率 $P(D)$，主観同時確率 $P(H,D)$ の回答分布を図 2 に示す．ベイズ解 (図 2 の破線) は $P(H|D)$ が 55.0%，$P(D)$ が 90.0%，$P(H,D)$ が 49.5% である．実験参加者の回答の中央値と中央値の 95% 信頼区間 (図 2 の網掛け範囲) は，それぞれ $P(H|D)$ が 50.0% (45.0–52.5)，$P(D)$ が 50.0% (50.0–75.0)，$P(H,D)$ が 49.1% (47.5–49.3) となった (信頼区間の算出法を付録 C.1 に示す)．

$P(H|D)$ のベイズ解は，回答の 95% 信頼区間の 45.0–52.5 に含まれておらず，本課題に正答できているとはいい難い．$P(D)$ の回答分布のピークは 2 つあり，中央値はベイズ解から逸脱している．1 つのピークは基準率の値であるが，これは偏確率率構造を認識できていない参加者が多いためと考えられる．もう 1 つのピークは偏確率構造を認識できているが，ベイズ解に近い回答となっている．$P(H,D)$ はほとんどの回答が 1 点に集中しており，ベイズ的回答といえる．

$P(H|D)$ については 45 以上 65 以下の回答を，$P(D)$ については 70 以上の回答を正答とみなして両者の連関を見たところ，Est-D 課題に正答したもののうち BR 課題に正答したものは 67% (6/9)，Est-D 課題に正答しなかったもののうち BR 課題に正答したものが 90% (9/10) となり有意ではなかった ($\chi^2(1, N = 19) = 0.47$, $p = .50$)．

BR 課題の $P(H|D)$ の回答は，ベイズ解 (55.0)

---

3) Est-D/HD 課題の目的は課題確率構造の内的表象，つまり $P(H)$ と $P(D)$ や $P(H,D)$ との関係がどのように表象されているかを検討することである．したがって，$P(D)$ や $P(H,D)$ の確率値を数値で答えさせる方法は心的表象をうまく引き出すためにはあまり良い方法であるとは言えない．そのため，$P(H)$ と $P(D)$ や $P(H,D)$ の心的表象を引き出すために，バーの長さで回答させるという推定法を用いた．

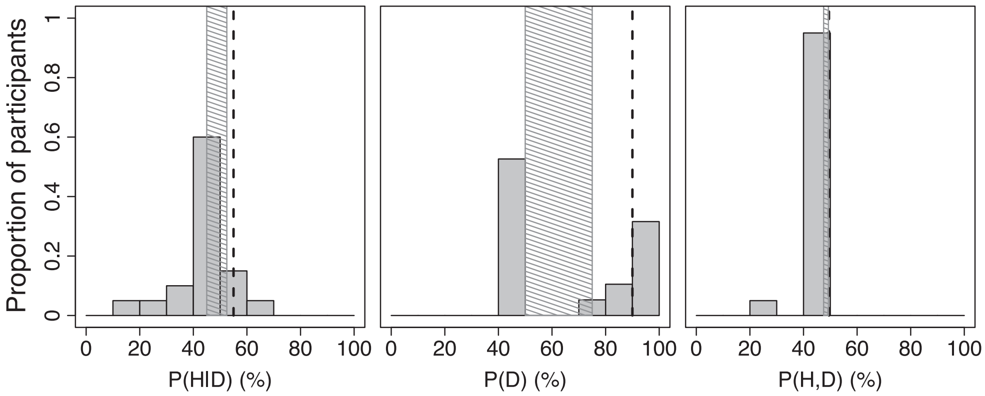


図 2 実験 1 の BR 課題，Est-D 課題，Est-HD 課題の回答分布．
図中の破線はベイズ解，網掛けは 95% 信頼区間を表す．

が回答の 95% 信頼区間の 45.0–52.5 に含まれていなかったが，参加者は単一事象が生起する確率を聞かれたときに「50」と答えがちである (Bruine de Bruin, Fischhoff, Millstein, & Halpern-Felsher, 2000) ことを考えれば，中央値の 50.0 をベイズ解に近いと見ることも可能である．したがって，本実験結果のみをもって基準率無視説が否定されたとするのは，少々強引な結論という印象を与えるかもしれない．そこで，無視説の妥当性について，実験 1 とは別の観点から再検討するのが適切と考え，次の実験を行うことにした．

3. 実験 2：基準率が大きいとき

Tversky & Kahneman (1980) の基準率無視説によれば，課題の基準率が大きければ，基準率が無視されても正解に近い値を答えられるはずである．なぜなら，式 (2) において，尤度オッズと基準率オッズの両方が高ければ，互いに打ち消しあうことはないため，尤度オッズの高さだけを考慮して高い事後確率を答えても，むしろ正解に近くなるからである．

それに対して，等確率性仮説は，基準率の高低に関わらず課題構造の認識が正答の鍵であると考える．よって，標準的な基準率課題（基準率が非常に小さい）と対照的な基準率が高い変形課題でも，課題に偏確率構造が存在する限り，正答は難しいことを予測する．実験 2 では，この予測が正しいかどうかを確認した．

3.1 方 法

3.1.1 実験参加者

大学生 20 名が実験に参加した．実験は集団により行われた．

3.1.2 課題と手続き

基準率が大きくなるよう $P(H) = .8$ かつ偏確率構造を保つよう $P(D) = .09$ とし課題を作成した．実験 2 で用いた確率構造を図 1-IV に示す．ストーリーは確率値が不自然にならないよう考慮し，以下を用いた．

> あなたは病院で働く医者であると仮定してください．近年，X 症候群という難病の感染例が報告されるようになりました．あなたに求められているのは次の情報から，患者が X 症候群を再発しているかどうかを判断することです．
>
> X 症候群の再発率は 80% です．X 症候群を再発しているならば，高血圧である確率は 10% です．X 症候群を再発していないとしても，高血圧である確率は 5% です．X 症候群の経験がある患者の血圧を測ったところ高血圧でした．この患者が X 症候群を再発している確率は何 % でしょうか？　直感で答えてください．
>
> (正解： 約 88.9 %)

実験 1 と同様の手続きで，BR 課題，Est-D 課題，Est-HD 課題を実施した．

3.2 結果と考察

実験 2 で得られた主観事後確率 $P(H|D)$，主観周

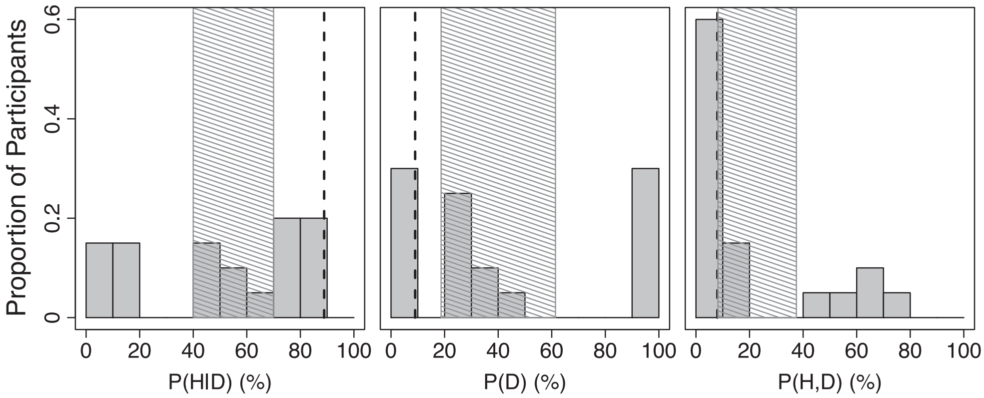


図 3 実験 2 の BR 課題，Est-D 課題，Est-HD 課題の回答分布．
図中の破線はベイズ解，網掛けは 95% 信頼区間を表す．

辺確率 $P(D)$，主観同時確率 $P(H,D)$ の回答分布を図 3 に示す．ベイズ解 (図 3 の破線) は，$P(H|D)$ が 89.0%，$P(D)$ が 9.0%，$P(H,D)$ が 8.0% である．参加者の回答の中央値と中央値の 95% 信頼区間 (図 3 の網掛け範囲) は，それぞれ，$P(H|D)$ が 57.5% (40.0–70.0)，$P(D)$ が 27.2% (18.7–61.4)，$P(H,D)$ が 10.0% (8.2–37.5) となった．

$P(H|D)$ の回答はほぼ一様に分布しており，ベイズ解が 95% 信頼区間に含まれていない．正答できている参加者はほとんどいないことがわかる．$P(D)$ の回答も，ベイズ解が 95% 信頼区間に含まれておらず，ベイズ解から逸脱した回答が多い．$P(H,D)$ の回答はベイズ解に近いといえるだろう．

結論として，実験 1, 2 の結果は基準率無視説の予測と一致しなかった．基準率を「無視」することが正答に導く課題（実験 1）でも，尤度オッズと基準率オッズが打ち消しあうことがないために，基準率を「無視」しても十分正答に近い回答が得られるべき課題（実験 2）でも，無視説が予測する高い正答率は得られなかった．それに対して，偏確率構造を持つ両実験の結果は，等確率性仮説の予測に一致するものであった．

4. 実験 3：偏確率構造認識の手段としての自然頻度形式

Gigerenzer & Hoffrage (1995) は，自然頻度表現による促進効果を計算しやすさによって説明した．基準率課題はベイズの定理を用いて規範解が得られるが，課題が確率で表現された場合は，ベイズの定理に当てはめる数値を自分で計算しなくてはならないため複雑である．一方，自然頻度で表現された場合には，計算に必要な数値を課題文から直接使って単純な算術計算で解を導くことができる．この考えは，進化的観点，すなわち，「心は確率論が整備されるずっと以前から人が用いてきた情報形式である自然頻度形式に調節されている[4)]」という考え (Gigerenzer & Hoffrage, 1995, p. 697) と整合的である．

しかし，Hattori & Nishida (2009) により，確率のみで情報を提示しても劇的に正答率が上昇することが示されたことから，情報の表現形式自体が決定的な要因ではないことがわかった．つまり，人が基準率課題に正答できるか否かは情報の表現によらない．等確率性仮説の考えでは，課題構造が認識しやすいか否かによって課題の難易度が変わる．

それでは，自然頻度による促進効果も等確率性仮説によって説明されるだろうか．Hattori & Nishida (2009) の 2 つの実験と本稿の実験 1, 2 の結果は，等確率性仮説を支持し，他の説明を否定するものであった．しかし，等確率性仮説が自然頻度による促進効果も説明可能かどうかについては，まだ検討されていない．自然頻度形式による情報の提示は，偏確率構造の認識を促すと言えるだろうか．自然頻度の促進効果は，偏確率構造認識の促進効果と同じものだと言えるだろうか．

本実験では，情報の提示形式と課題構造の認識のしやすさを操作して実験を行い，自然頻度による促

4) 原文では “… the mind is tuned to frequency formats, which is the information format humans encountered long before the advent of probability theory.”

進効果が，偏確率構造認識による促進効果と同じものと考えられるかどうかについて，正答率と内的表象の両面から検討した．もし，最終出力としての正答率上昇の程度に関して両者間に差がなく，また，内的表象に関しても差異が認めにくいとすれば，自然頻度形式による促進効果は，この情報提示形式が偏確率構造の認識を促すために生じると考えることが可能であろう．

4.1 方　法

4.1.1 実験計画と実験参加者

2 (提示形式：自然頻度 vs. 確率) × 2 (課題構造の認識：検査 vs. 咳) の計 4 条件を設け，被験者間計画を用いた．大学生 65 名が集団で実験に参加した．実験参加者は検査・確率条件 ($n = 16$)，検査・頻度条件 ($n = 16$)，咳・確率条件 ($n = 13$)，咳・頻度条件 ($n = 20$) にランダムに配置された．

4.1.2 課題と手続き

Gigerenzer & Hoffrage (1995) のカバーストーリーを変更したものを，咳・頻度条件で使用した．確率提示条件の課題は，課題中の数値情報を百分率で提示したもので，Hattori & Nishida (2009) の実験 1 の咳条件で使用されたものと同じ課題である．咳・頻度条件で用いた課題文を以下に示す．

> 近年，X 症候群という難病の感染例が報告されるようになりました．この X 症候群は風邪と似た症状が出ることが知られています．
>
> さて，あなたは病院で働く医者であると仮定してください．あなたに求められているのは次の情報から患者が X 症候群に感染しているかどうかを判断することです．
>
> 日本人の 1000 人に 10 人はX症候群に感染しています．X 症候群に感染している 10 人のうち 8 人は，咳の症状があります．X 症候群に感染していない 990 人のうち 95 人も，咳の症状があります．咳をしている患者の数を想像してください．本当に X 症候群である患者は何人でしょうか？直感で答えてください．
>
> ＿＿＿ 人中 ＿＿＿ 人．
>
> (正解： 103 人中 8 人 (約 7.8 %))

咳・確率条件，検査・頻度条件，検査・確率条件で用いた課題は付録 A に示す．実験 3 で用いた確率構造を図 1-I に示す．この課題では，咳が珍しい症状ではないという一般知識が，偏確率構造 $P(H) \ll P(D)$ の認識を促進する．検査条件では，「咳の症状がある」を「検査で陽性となる」とした．実験 1 と同様の手続きで，BR 課題，Est-D 課題，Est-HD 課題を実施した．

4.2 結果と考察

4.2.1 正答率改善の程度

BR 課題のベイズ解は 7.8% である．中央値と最頻値はそれぞれ，検査・確率条件 (60.0 / 60.0)，検査・頻度条件 (5.0 / 1.0)，咳・確率条件 (1.0 / 1.0)，咳・頻度条件 (8.4 / 10.0) であった．図 4 に各条件ごとの $P(H|D)$ の回答分布を示す．Kruskal-Wallis 検定を行ったところ，中央値間に有意な差が見られたので ($\chi^2(3) = 16.69$, $p < .01$)，続いて Scheffe の方法による対比較を行った．検査・確率条件と検査・頻度条件間 ($\chi^2(3) = 10.75$, $p < .05$)，検査・確率条件と咳・確率条件間 ($\chi^2(3) = 13.59$, $p < .01$) に有意な差が見られた．検査・確率条件と咳・頻度条件間の差は有意ではなかった ($\chi^2(3) = 6.81$, $p = .08$)．他の条件間では有意な差は見られなかった．検査・確率条件以外の 3 つの条件では，実験参加者の回答がベイズ解に近づいており，基準率を使用して課題に回答したことがわかる．

回答分布とベイズ解の値 (7.8%) から判断し，10 以下の回答を正答，70 以上の回答を典型的誤答，10 以上 70 未満の回答をその他とみなし，参加者の回答比率を求めた (図 5)．正答率は，検査・確率条件が 25%，検査・頻度条件が 75%，咳・確率条件が 77%，咳・頻度条件が 80% であった．正答率に有意な差が見られたため ($\chi^2(6, N = 65) = 23.26$, $p < .01$) 残差分析を行ったところ，検査・確率条件のみ有意に正答率が低かった ($p < .05$)．

以上により，カバーストーリーを変更して偏確率構造を認識させることにより，正答率が上昇するという Hattori & Nishida (2009) の結果と，自然頻度を用いれば正答率が上昇するという Gigerenzer & Hoffrage (1995) の結果を再確認できた．さらに，検査・確率条件以外の 3 つの条件では中央値，正答率ともに有意な差は見られなかったことから，偏確率構造の認識を促進させた場合と自然頻度を用いた場合とで，同程度の上昇効果を得られることが示された．

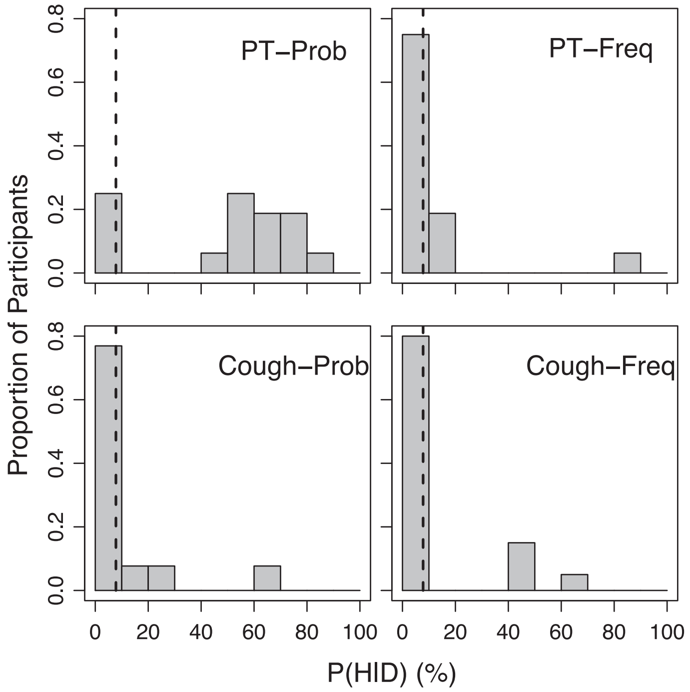


図 4　実験 3 の BR 課題の回答分布．PT-Prob, PT-Freq, Cough-Prob, Cough-Freq はそれぞれ，検査・確率，検査・頻度，咳・確率，咳・頻度条件を表す．図中の破線はベイズ解を表す．

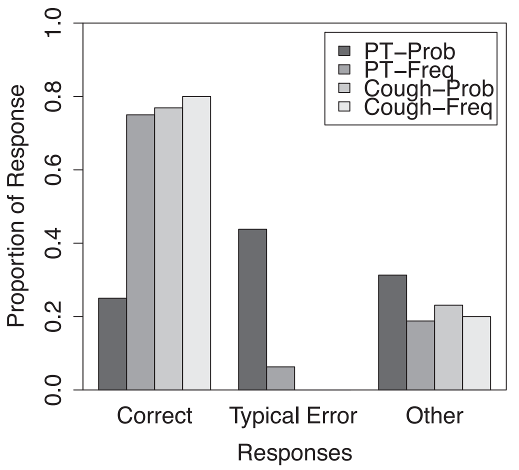


図 5　実験 3 の BR 課題のカテゴリ別回答分布を条件別に表したもの．条件名は図 4 と同じ

4.2.2　偏確率構造の内的表象

実験参加者が課題の確率構造に関してどのような内的表象を保持しているか検討するため，2 つの確率推定課題の分析を行った．

Est-D 課題　この課題のベイズ解は 10.3% である．図 6 に各条件ごとの $P(D)$ の回答分布を示す．得られた $P(D)$ の中央値は，検査・確率条件が 1.22，検査・頻度条件が 0.89，咳・確率条件が 1.28，咳・頻度条件が 6.85 であった．

Est-D 課題では，課題の等確率性の認識に成功した参加者と，失敗した参加者の 2 群が存在すると考えられる．つまり質的に異なる 2 つの分布が混合された双峰性の回答分布が得られると予想される．そこで，まず各条件ごとにカーネル密度推定法によって確率密度を求め，どのような分布形状をしているかを検討した (算出法を付録 C.2 に示す)．カーネル密度推定を行うためにはパラメータであるバンド幅 h を指定する必要があるが，本分析ではすべての条件で同じバンド幅を持つと仮定し，すべての条件をまとめてカーネル密度推定法を行い，Silverman (1986) の「経験則」によって計算されたバンド幅 $h = 1.666$ を，各条件における密度推定にも用いた．カーネル関数にはガウスカーネルを用いた．検査・確率条件はピークを 1 つ持つ単峰性の分布であるのに対し，検査・頻度，咳・確率，咳・頻度条件は，ピークを 2 つ持つ双峰性の分布を持つことが確認された (図 6).

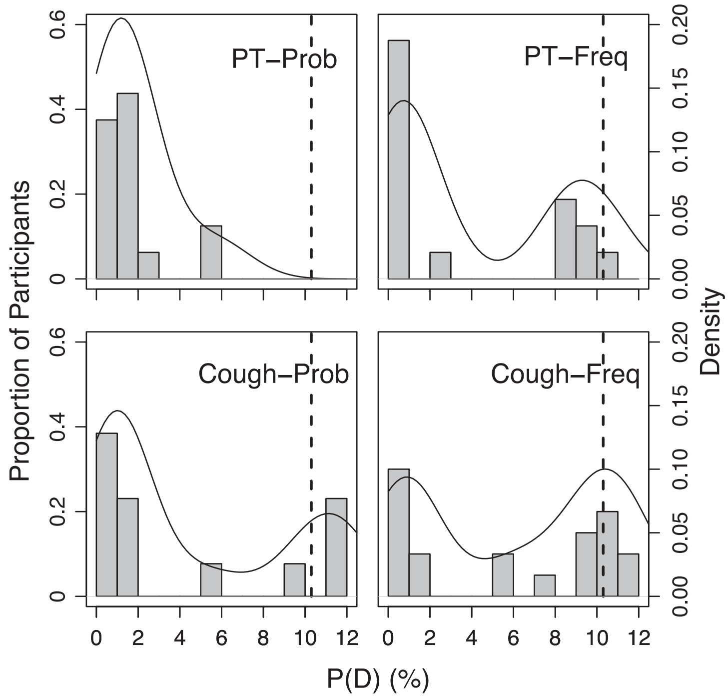


図 6 実験 3 の Est-D 課題の回答分布とカーネル密度推定法により推定された確率密度．パラメータであるバンド幅にはすべての条件で 1.666 の値を用いた（本文参照）．条件名は図 4 と同じ．図中の破線はベイズ解を表す．

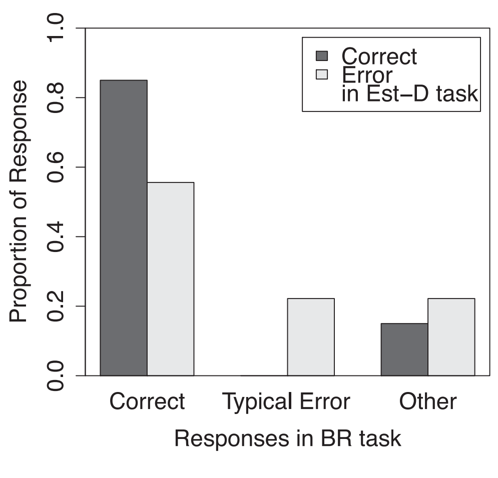


図 7 実験 3 の BR 課題のカテゴリ別回答分布を Est-D 課題の正誤別に表したもの．回答カテゴリは，正答，典型的誤答，その他の誤答を表す．

さらに，各条件の分布の形状に差があるか否かを検討するため，カルバック・ライブラ (KL) 情報量 (Kullback & Leibler, 1951) を用いた．確率分布 P から Q の KL 情報量は次のように定義される．

表 1 実験 3 の Est-D 課題で得られた回答分布の条件間の KL 情報量．

P \ Q	検査・確率	検査・頻度	咳・確率	咳・頻度
検査・確率	0.00	0.44	0.39	0.73
検査・頻度	1.17	0.00	0.11	0.10
咳・確率	1.38	0.11	0.00	0.12
咳・頻度	2.09	0.11	0.12	0.00

$$D_{KL}(P||Q) = \int p(x) \log \frac{p(x)}{q(x)} dx \quad (3)$$

ここで $p(x)$ と $q(x)$ は，それぞれ確率分布 P と Q の密度である．KL 情報量は 2 つの確率分布 (P, Q) の差の尺度であり，常に負でない値をとる ($D_{KL}(P||Q) \geq 0$)．KL 情報量は対称性を持たない ($D_{KL}(P||Q) \not\equiv D_{KL}(Q||P)$) ため，一般的な意味での距離とは異なるが，確率分布の不一致度を示す指標となる．2 つの確率分布が完全に一致するときに 0 の値をとり，確率分布がずれているほど値が大

きくなる．

カーネル密度推定により推定された密度を用いて KL 情報量を計算した．また，密度の値が 10^{-4} 未満の場合は計算のため 10^{-4} を用いた．行を P，列を Q としたときの各条件間の KL 情報量を表 1 に示す．表 1 より，検査・確率 条件のみが他の条件に対して大きな値をとっていることがわかる．また，検査・頻度，咳・確率，咳・頻度条件は，すべて 0 に近い値をとっていることから，この 3 条件の確率分布は似通った分布形状を有していると考えられる．すなわち，検査・確率条件のみが他条件と比べ異なる分布を有しているといえる．これは，$P(D)$ に関して，検査・確率条件のみが異なる内的表象を保持していることが示唆される．

次に BR 課題と Est-D 課題との関係を検討した．Est-D 課題の回答を回答分布とベイズ解の値 (10.3%) から判断して 7.0 以上の回答を正答，7.0 未満の回答を誤答とみなし，条件をプールして BR 課題の正答率との関連を見た (図 7)．BR 課題の正答率に有意な差が見られたので ($\chi^2(2, N = 65) = 6.66$, $p < .05$)，続いて残差分析を行った．Est-D 課題での正答群は BR 課題での正答率が高く，典型的誤答が低かった ($p < .05$)．

以上の結果は，$P(D)$ の値を正しく認識していると，$P(H|D)$ を正しく答えることができることを示している．これは，$P(D)$ を明示的に示した場合には正答率が上昇したという Gigerenzer & Hoffrage (1995) の結果に一致する．しかし，Gigerenzer & Hoffrage (1995) はこの結果を，計算が簡単になったためと解釈したが，$P(D)$ を明示的に示さなくとも正答率が上昇した今回の結果より，課題の確率構造を正しく認識できたためと解釈する方が妥当だと考えられる．

Est-HD 課題　この課題のベイズ解は 0.8% である．図 8 に各条件ごとの $P(H, D)$ の回答分布を示す．得られた $P(H, D)$ の中央値は検査・確率条件が 0.82，検査・頻度条件が 0.81，咳・確率条件が 0.79，咳・頻度条件が 0.82 であった．Kruskal-Wallis 検定を行ったところ条件間に有意な差はなかった ($\chi^2(3) = 2.73$, $p = .43$)．

次に BR 課題の正答と $P(H, D)$ 推定課題の正答との関係を検討した．0.8 ± 0.1 を正答とみなしたところ，Est-HD 課題に正答したもののうち BR 課題にも正答できたものは 72% (34/47)，Est-HD 課題に正答しなかったもののうち BR 課題に正答できたものは 44% (8/18) で，有意差は見られなかった ($\chi^2(1, N = 65) = 3.29$, $p = .07$)．

以上のことから $P(H, D)$ に関して，すべての条件でベイズ的な内的表象を保持していることが示唆される．2 つの事象が同時に起こる確率は，2 事象の関係を把握するうえで特に重要である．そのため，偏確率構造の認識しやすさに関わらず正しく $P(H, D)$ を推定できたと考えられる．

3 つの課題の結果から，次の 2 つのことが示された．まず，BR 課題において，咳を用いた条件（頻度と確率）と自然頻度を用いた条件（咳と検査）が，極めて類似した回答分布（最終出力）を示した．さらに，これらの条件は，Est-D/HD 課題の結果についても類似した結果を示した．この結果が意味することは，最終的な出力の観点からも内的表象の観点からも，自然頻度形式による情報提示の効果は，代替原因の示唆による偏確率構造の認識の効果と区別がつかないということである．このことから，自然頻度表現の正答率上昇は，計算が簡単になるという表面的な理由によるものではなく，自然頻度表現が偏確率構造の認識を促進させるためという解釈ができる．すなわち，等確率性仮説は，自然頻度仮説を包含する上位概念であると言えるだろう．

5. 総 合 考 察

本研究は，基準率無視説 (Tversky & Kahneman, 1980) と自然頻度仮説 (Gigerenzer & Hoffrage, 1995) を批判的に検討した．実験 1, 2 の結果より無視説は否定され，実験 3 の結果より，自然頻度による促進効果が課題構造の認識による可能性が示唆された．

そこで本節では，無視説の理論的基盤とされてきた代表性ヒューリスティックと等確率性の関係，また，等確率性に近い考えである反転錯誤の位置づけ，さらには，課題構造認識を促す方法としての自然頻度の位置づけなどについて考察する．

5.1　基準率無視と代表性ヒューリスティック

基準率無視の理論的背景として，代表性ヒューリスティック (Tversky & Kahneman, 1974) が考えられてきた．ここでは，基準率無視と代表性ヒューリスティックの整合性について考察したい．代表性ヒューリスティックによる判断とは，$P(H|D)$ が D

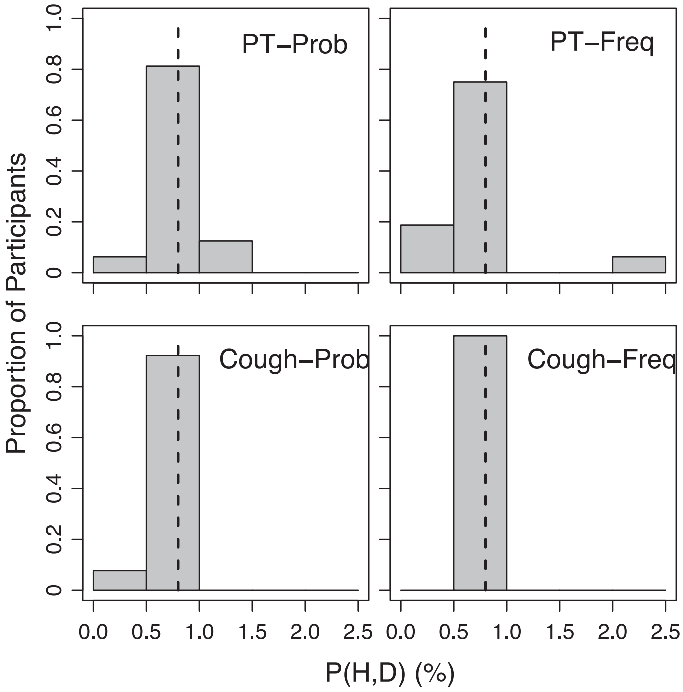


図 8 実験 3 の Est-HD 課題の回答分布.
条件名は図 4 と同じ．図中の破線はベイズ解を表す.

と H の類似性によって推論されることを意味する．しかし，この類似性がどのようにして計算され，また類似性からどのようにして確率値へ変換されるのか明確にされておらず，代表性ヒューリスティックの概念は曖昧でよく定義されていない (Shanteau, 1989) と批判されてきた.

Gigerenzer & Murray (1987) は，代表性ヒューリスティックの概念を尤度 $P(D|H)$ によって定義することを提案した．つまり，代表性ヒューリスティックによる判断とは尤度 $P(D|H)$ による判断と言い換えることができる．類似性を尤度で置き換えた代表性ヒューリスティックによる判断は Gigerenzer & Hoffrage (1995) のフィッシャー的アルゴリズムの特別ケースである．フィッシャー的アルゴリズムとは，基準率課題において尤度 $P(D|H)$ の情報のみを用いて，事後確率 $P(H|D)$ を推定する認知的アルゴリズム $P(H|D) \simeq P(D|H)$ である (Gigerenzer & Hoffrage, 1995).

しかし，この定義によると，代表性ヒューリスティックと基準率無視説は整合しない．無視説によれば，$P(H) = P(\bar{H})$ ゆえ，式 (1) は $P(H|D) = P(D|H)/[P(D|H)+P(D|\bar{H})]$ となるため，冒頭に示した乳がん課題 (Eddy, 1982) の推定値は 89.3% となる．しかし，代表性ヒューリスティックによる推定値は $P(H|D) = P(D|H) = 80.0\%$ である．無視説と代表性ヒューリスティックが一致するのは，$P(D|H) + P(D|\bar{H}) = 1$ が満たされているときに限られることがわかる．実は，Kahneman らが最初に示したタクシー課題 (Tversky & Kahneman, 1982) においては，たまたまこの条件が満たされているが，このような条件を満たす課題の方がむしろ特殊である[5]．よって，代表性ヒューリスティックを無視説の理論的背景と考えるのは難しい.

さらに，Gigerenzer & Hoffrage (1995) は，$P(H)/P(D) = 1$ であればベイズの定理はフィッシャー的アルゴリズムと同じ結果を導くと述べている (p. 698)．この仮定は，等確率性の仮定 $P(H) = P(D)$ と同じであり，尤度により類似性を定義した代表性ヒューリスティックは等確率性仮説と同一とみなすことができる．つまり，基準率無視を説明すると考えられてきた代表性ヒューリスティックは，Gigerenzer らの定義を採用する限りに

5) Gigerenzer & Hoffrage (1995) が実験 1 で用いた 15 個の課題のうち，この条件を満たしているものは 2 個のみである.

おいては，基準率無視説よりむしろ等確率性仮説と整合的であると言えよう．

5.2　反転錯誤と一般的知識

人は与えられた条件つき確率 $P(D|H)$ を $P(H|D)$ と混同してしまうという考えもある．これは反転錯誤 (inverse fallacy) と呼ばれている (Villejoubert & Mandel, 2002)．反転錯誤も等確率性仮説と近い考え方であり，予測結果の多くを共有する．しかし，なぜ反転錯誤が生じるのかが説明されていない．この考え方では回答者は $P(H|D)$ をたずねられたとき，常に $P(D|H)$ を答えることになり常に間違えた回答をすることになる．一方，等確率性仮説は $P(H) \simeq P(D)$ を仮定する場合に反転錯誤が生じると考える．

Krynski & Tenembaum (2007) は次のような例を出し，反転錯誤による説明の限界を示している．「飛行機事故のときに死ぬ確率はほぼ 100% である」(P(death|plane crash) $\simeq 1.0$) とする．この確率の条件つきを逆にすると，「死因のほとんどは飛行機事故である」(P(plane crash|death) $\simeq 1.0$) になってしまう．ここから Krynski & Tenembaum (2007) は，反転錯誤は 2 つの条件つき確率が同程度である ($P(D|H) \simeq P(H|D)$) と回答者が期待する場合しか起こりえないとしている．

この例は，一般的な知識が正しい確率的構造を認識させている例といえるだろう．Hattori & Nishida (2009) の実験 1 では，本研究の実験 3 で用いたような一般知識が活用されるカバーストーリーにおいては反転錯誤が起こらず，ベイズ的な解にたどり着けることが示された．これは，偏確率構造を認識できれば反転錯誤が生じにくいことを示している．つまり，反転錯誤が生じるのは，回答者が等確率性を仮定しているからだといえよう．

5.3　自然頻度と課題構造

一般的な知識を用いるのと同様に，自然頻度による情報提示は，課題構造を認識させる方法であることが実験 3 で示された．自然頻度による提示が正答率を上昇させる効果があることは疑う余地がないが，なぜ正答率が上昇するかについて，理論的な説明があまりなされていない．Gigerenzer & Hoffrage (1995) による進化論に依拠した説明，すなわち計算が簡単になるという説明は，顕在的プロセスに重点を置くものであり，基準率課題の難しさの本質を捉えているとは思われない．

等確率性仮説では，課題構造とその心的表象に焦点を当て，課題構造と心的表象が乖離している場合には難しい課題になり，一致している場合には容易な課題になると予測する．実験の結果はこの予測に一致するもので，自然頻度形式の情報提示に独特の効果を見出すことができなかった．

課題構造の認識を促す方法は複数あると考えられる．1 つ目は Hattori & Nishida (2009) や本研究でも用いられた一般知識の利用である．2 つ目は今回検証された自然頻度による情報提示である．おそらく，この他にも課題構造の認識させる方法があると考えられる．たとえば，直接経験によって実際に事例を体験すると，基準率錯誤が起こりにくいことが知られているが (e.g., Christensen-Szalanski & Beach, 1982)，この促進効果も，課題構造の認識が生んだものである可能性がある．これらの方法は，デフォルトの等確率性を仮定した心的表象から，課題構造に一致した心的表象を再構成することを促す効果があると考えられる．

6.　まとめと展望

本研究の結果は，基準率課題での誤答は基準率の無視によって生じるのではなく，回答者が 2 事象間の等確率性を仮定しているために生じているという Hattori & Nishida (2009) を支持する．実験の結果から，(1) 基準率無視説による説明には限界があり，等確率性仮説でなければ説明できない現象があることが示され，(2) 自然頻度仮説は確率構造の認識を促進させる一手法であり，等確率性仮説の下位概念に位置づけられることが示唆された．

等確率性の仮定は，人の思考全般と深く関連していると考えられる．条件文推論 (Hattori, 2002)，因果帰納 (Hattori & Oaksford, 2007)，仮説検証 (田村ら, 2010) といった思考において，人が等確率性を仮定していることが示されている．近年では，服部 (2008) によって，人の多くの思考過程が等確率性仮定の観点から統一的に説明可能であることがまとめられている．このように，人が等確率性を仮定しているという証拠は累積的に集まりつつある．この次に問題となるのは，なぜ等確率性を仮定するかについての包括的で整合的な説明であろう．

人の認知システムは環境に最適化されている (Anderson, 1990) という考えからすれば，等確率性仮定には適応的合理性があると考えられる．Tversky & Kahneman (1983) は，会話で獲得できる情報量の期待値の観点から代表性ヒューリスティックについて論じた．不確実性状況下におけるメッセージ伝達は，メッセージの内容と確実性の両面から評価され，メッセージの価値は，内容の確からしさによって重みづけられた期待価値として定義される．Grice (1975) の協調の原則に基づく会話において，価値の高いメッセージを選択するために代表性ヒューリスティックが助けになっていると考えることができる．もし，代表性ヒューリスティックが等確率性仮説と同一視できる (§5.1) ならば，等確率性の仮定が，語用論的な意味での情報伝達の最適化に寄与しており，言語を含む認知環境において適応的合理性を有している可能性が同じ議論から示唆されるだろう．

今後は，このような観点からの実証的研究に加えて，同じく人が持つ思考の傾向性と考えられる稀少性仮定 (Oaksford & Chater, 1994) と等確率性仮定の関連を明らかにすることや，思考のベースとなる概念の理論の中での等確率性仮定の位置づけを探ることなどが研究課題となるだろう．

謝　辞

本研究に対して，足立浩平教授，市川伸一教授，高野陽太郎教授（ABC 順）から有益なコメントを頂いた．また，2 名の匿名の査読者と中村國則担当編集委員より頂いたコメントは，原稿の改善に大きく貢献した．実験実施にあたっては，足立浩平教授にご尽力頂いた．以上，ここに記して謝意を表する．なお，本研究の実施と本稿の執筆は，日本学術振興会科学研究費補助金 22500247（研究代表者：服部雅史）の資金援助を受けた．

文　献

Anderson, J. R. (1990). *The Adaptive character of thought.* Hillsdale, NJ: Erlbaum.

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*, **30**, 241–254.

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica*, **44**, 211–233.

Bishop, C. M. (2006). *Pattern recognition and machine learning.* New York: Springer.

Bruine de Bruin, W., Fischhoff, B., Millstein, S. G., & Halpern-Felsher, B. L. (2000). Verbal and numerical expressions of probability: "It's a fifty-fifty chance." *Organizational Behavior and Human Decision Processes*, **81**, 115–131.

Campbell, M. J., & Gardner, M. J. (1988). Calculating confidence intervals for some non-parametric analyses. *British Medical Journal*, **296**, 1454–1456.

Casscells, W., Schoenberger, A., & Grayboys, T. (1978). Interpretation by physicians of clinical laboratpry results. *New England Journal of Medicine*, **299**, 999–1000.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, **58**, 1–73.

Christensen-Szalanski, J. J. J., & Beach, L. R. (1982). Experience and the base-rate fallacy. *Organizational Behavior and Human Performance*, **29**, 270–278.

Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*, 249–267. Cambridge: Cambridge University Press.

Evans, J. St. B. T., Handley, S. J., Perham, N., Over, D. E., & Thompson, V. A. (2000). Frequency versus probability formats in statistical word problems. *Cognition*, **77**, 197–213.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve bayesian reasoning without instruction: Frequency formats. *Psychological Review*, **102**, 684–704.

Gigerenzer, G., & Murray, U. (1987). *Cognition as intuitive statistics.* Hillsdale, NJ: Erlbaum.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J.L. Morgan (Eds.), *Syntax and semantics 3: Speech acts*, 41–58. London: Academic Press.

Hattori, M. (2002). A quantitative model of optimal data selection in Wason's selection task. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, **55A**, 1241–1272.

服部 雅史 (2008). 推論と判断の等確率性仮説：思考の対称性とその適応的意味. 『認知科学』, **15**, 408–427.

Hattori, M., & Nishida, Y. (2009). Why does the base rate appear to be ignored? The equiprobability hypothesis. *Psychonomic Bulletin & Review*, **16**, 1065–1070.

Hattori, M., & Oaksford, M. (2007). Adaptive non-interventional heuristics for covariation detection in causal induction: Model comparison and rational analysis. *Cognitive Science*, **31**, 765–814.

Hollander, M., & Wolfe, D. A. (1973). *Nonparametric statistical methods*. New York: John Wiley & Sons.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, **22**, 79–86.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Normative, descriptive and methodological challenges. *Behavioral and Brain Sciences*, **19**, 1–53.

Krynski, T. R., & Tenembaum, J. B. (2007). The role of causality in judgement under unertainty. *Journal of Experimental Psychology: Genaral*, **136**, 430–450.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, **101**, 608–631.

Shanteau, J. (1989). Cognitive heirustic and biases in behavioral auditing: Review, comments and observations. *Accounting Organizations and Society*, **14**, 165–177.

Silverman, B. W. (1986) *Density Estimation*. London: Chapman and Hall.

Sloman, S. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes*, **91**, 296–309.

田村 昌彦・服部 雅史・三輪 和久 (2010). 仮説検証過程における確信度更新と検証系列：情報獲得モデルによる検討. 『認知科学』, **17**, 180–195.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, **185**, 1124–1130.

Tversky, A., & Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in social Psychology, Vol.1*, 49–72. Hillsdale, NJ: Lawrence Erlbaum Associates.

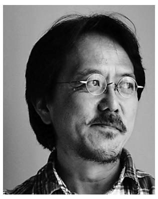
Tversky, A., & Kahneman, D. (1982). Evidental impact of base rate. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*, 153–160. Cambridge, UK: Cambridge University Press.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallcy in probability judgment. *Psychological Review*, **90**, 293–315.

Villejoubert, G., & Mandel, D. R. (2002). The inverse fallacy: An account of deviations from Bayes's Theorem and the additivity principle. *Memory & Cognition*, **30**, 171–178.

(Received　15　June　2010)
(Accepted　22　Nov.　2010)

西田　豊 (学生会員)
2006 年立命館大学文学部卒業. 2008 年大阪大学大学院人間科学研究科博士前期課程修了. 現在, 同研究科博士後期課程在学. 関心は推論, 判断, 概念など高次認知機能の適応的合理性. それらの数理モデル化や方法論の開発にも興味を持つ. 日本心理学会, 日本行動計量学会, 各会員.



服部　雅史 (正会員)
1996 年北海道大学大学院文学研究科博士後期課程単位取得退学. 博士（文学）. 1997 年より立命館大学文学部. 2003–2004 年英国（ウェールズ）カーディフ大学心理学部客員研究員. 現在, 立命館大学文学部教授・米国ブラウン大学認知言語心理科学部客員研究員. 推論, 意思決定, 問題解決などの高次認知機能の研究に従事. 人間の思考・言語・コミュニケーションの適応的合理性に興味を持つ. 日本心理学会（2007– 編集委員）, 日本基礎心理学会, Cognitive Science Society, Psychonomic Society ほか会員.
`hat@lt.ritsumei.ac.jp`

付　録

A.　BR 課題

実験 3 で用いた課題文を以下に示す．各条件の課題文の前には以下の状況設定の文が挿入された．

近年，X 症候群という難病の感染例が報告されるようになりました．この X 症候群は風邪と似た症状が出ることが知られています．

さて，あなたは病院で働く医者であると仮定してください．あなたに求められているのは次の情報から患者が X 症候群に感染しているかどうかを判断することです．

A.1　咳・確率条件

X 症候群の感染率は 1% です．X 症候群に感染しているならば，咳の症状がある確率は 80% です．X 症候群に感染していないとしても，咳の症状がある確率は 9.6% です．今，患者が咳をしています．この患者が本当に X 症候群である確率は何 % でしょうか？ 直感で答えてください．

A.2　検査・頻度条件

日本人の 1000 人に 10 人は X 症候群に感染しています．X 症候群に感染している 10 人のうち 8 人は，検査で陽性となります．X 症候群に感染していない 990 人のうち 95 人も，検査で陽性となります．検査で陽性でとなった患者の数を想像してください．本当に X 症候群である患者は何人でしょうか？ 直感で答えてください．

A.3　検査・確率条件

X 症候群の感染率は 1% です．X 症候群に感染しているならば，検査で陽性となる確率は 80% です．X 症候群に感染していないとしても，検査で陽性となる確率は 9.6% です．今，患者の検査が陽性でした．この患者が本当に X 症候群である確率は何 % でしょうか？ 直感で答えてください．

B.　Est-D/HD 課題

実験 1, 2, 3 で共通に用いた Est-D/HD 課題を以下に示す．この課題は，課題の確率構造の表象を探るための課題である．どちらの課題も，図を使用した一種のマグニチュード推定法を用いている．実験 1 において，Est-D 課題では事象 H (新薬) である人数 $N(H)$ を基準として，事象 D (症状改善) である人数 $N(D)$ を推定させた．Est-HD 課題では，$N(H)$ を基準として，事象 H and D (新薬 and 症状改善) である人数 $N(H, D)$ を評定させた．分析には，$N(D)$，$N(H, D)$ の値を $N(H)$ の値で除したものに $P(H)$ を乗じたものを $P(D)$，$P(H, D)$ として用いた．このとき確率値が 1 を超えたものは 1 として扱った．

なお，次ページに示す回答用紙サンプルは実験 1 で用いたものであるが，他の実験でも同様の形式を用いた．実験 2 では H:「X 症候群」，D:「高血圧」，H and D:「X 症候群 and 高血圧」として，実験 3 では H:「X 症候群」，D:「陽性」もしくは「咳」，H and D:「X 症候群 and 陽性」もしくは「X 症候群 and 咳」として用いられた．

C.　分析で用いた統計手法について

中央値の信頼区間とカーネル密度推定法について，その算出法を示す．サンプルサイズを n，データを x_i $(i = 1, \cdots, n)$ とする．

C.1　中央値の信頼区間

Wilcoxon の符号付き順位統計量に基づいて，中央値の信頼区間を算出した (Hollander & Wolfe, 1973; Campbell & Gardner, 1988)．$(1-\alpha)100\%$ の信頼区間は以下の手順で求まる．

(1)　あらゆる 2 つのデータの組合せについての $n(n+1)/2$ 個の平均値 $M_{ij} = (x_i + x_j)/2$ を算出し，$M^{(1)} \leq \cdots \leq M^{([n(n+1)/2])}$ のように並べる．

(2)　Wilcoxon の符号付き順位統計量の $(\alpha/2)$ 100% の値 $W_{\alpha/2}$ を求める．

(3)　$(1-\alpha)100\%$ の信頼区間 (θ_L, θ_U) は

$$\theta_L = M^{(W_{\alpha/2})}, \quad \theta_U = M^{([n(n+1)/2]+1-W_{\alpha/2})}$$

によって与えられる．

すなわち，$n(n+1)/2$ 個の M の値の小さい方から $W_{\alpha/2}$ 番目と，大きい方から $W_{\alpha/2}$ 番目の値が信頼限界である．

C.2　カーネル密度推定

カーネル密度推定法はノンパラメトリックな密度推定法の 1 つである (Bishop, 2006)．データが

1 次元のとき，ガウスカーネルを用いた推定密度は

$$p(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{(2\pi h^2)^{1/2}}\exp\left\{-\frac{(x-x_i)^2}{2h^2}\right\}$$

で得られる．ここで h はバンド幅である．各データ点 x_i を平均，h^2 を分散とした正規分布をすべてのデータ i について合計し，正規化のためにサンプルサイズ n で割ったものである．

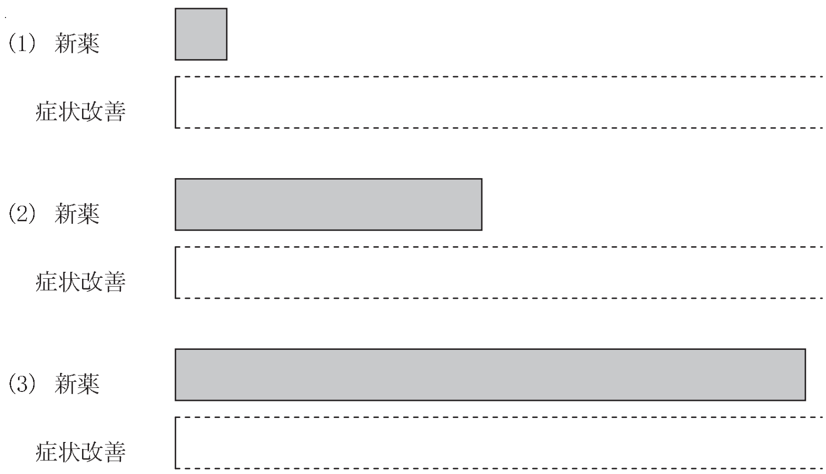
回答用紙サンプル

Est-D 課題

新薬を投与されている患者の数を模式的に図のように四角形の大きさで表します．症状が改善した患者の数を，四角形の大きさで表すとどのくらいの大きさになるでしょうか？ 直感で答えてください．下の点線の四角形は右端が開いていて，大きさが決まっていません．右端の縦線を 1 本書き入れて,「症状が改善した患者の数」を表す大きさの四角形を完成させてください.

答えは (1), (2), (3) の解答欄のうち，あなたが思い浮かべる「症状が改善した患者の数」を表す大きさの四角形を一番描きやすいものを 1 つだけ選んで書いてください．例えば，X 症候群の四角形より大きい四角形を描きたければ (1) に，X 症候群の四角形より小さい四角形を描きたければ (3) に，同じ程度の大きさなら (2) に描いてください.

(1), (2), (3) の四角形は大きさは違いますが，同じ人数を意味しています.



Est-HD 課題

新薬を投与されている患者の数を模式的に図のように四角形の大きさで表します．新薬を使って症状が改善した患者の数を，四角形の大きさで表すとするとどのくらいの大きさになるでしょうか？ 直感で答えてください．下の点線の四角形は右端が開いていて，大きさが決まっていません．右端の縦線を 1 本書き入れて,「新薬を使って症状が改善した患者の数」を表す大きさの四角形を完成させてください.