❍ 報告論題：–

『３囚人問題等に関する新ベイズ解』

❍ 報告者 (所属)・共著者 (所属)：–

任輝 (福岡大学)・奈良由美子 (放送大学)・藤本浩明 (福岡大学)

● 報告要旨：–

本稿では、３囚人問題およびモンティホール問題を取り扱う。通常、これらの解答は、ベイズの定理と確率樹形図を用いて恣意的に計算されがちである。便宜上ここでは、その計算結果を無理矢理ベイズ解と呼ぶことにする。しかし、この解を含め、「満足な解答は今のところない (松原, 2010, p.69)」。なぜならば、標本分布を明示しないばかりでなく、そのパラメータの事前分布として、「相性のよいお相手は決まっている (ハーバード学派)(松原, 2010, p.68)」にも関らず、自然共役分布を考慮しないからである。そこで、我々は、３点 (または３項) 分布を標本分布に、相性のよい自然共役的なディリクレ分布を事前分布に選ぶことによって、この種の問題の新ベイズ解を考案する。

■ 解説 ■

■§1. 問題の所在について：–

今野ら (2010, p.90) によると、間違いやすい確率・統計の問題があり、それは、条件つき確率の計算・推定時に生じやすい。なぜならば、確率の演繹的計算では、

❶ 同時分布から周辺分布を導出するだけで済む場合 (今野ら, 2010, p.92) と

❷ 事象の発生時間を遡るべく、ベイズの定理を用いる場合 (今野ら, 2010, p.94)

とがあるからで、また、統計量の帰納的推定でも、

❸ 標本はある分布に従うが、パラメータは未知だけど一定とおくだけの場合と

❹ 標本同様、パラメータもある分布に従うため、ベイズの定理が必要な場合

とがあるからである。つまり、❶ と ❸ により、問題が何であろうが、ベイズの定理さえ用いれば良いと言うわけではない。そこで本稿は、３囚人問題およびモンティホール問題を ❶ に属する問題として捉えながら、統計量を推定する際には、❹ の手法を用いる。すなわち、我々は、標本の同時分布として３点 (または３項) 分布、パラメータの事前分布としてディリクレ分布を選び、事後分布を計算する際に、ベイズの定理を使用する。なお、標本分布の周辺分布は２点 (または２項) 分布に、ディリクレ分布はベータ分布に退化することが、それぞれ、予想される。

■§2. ベイズの定理について：–

トーマスベイズの原典では、各々同様に確からしい、３項 ($n = 1$ は３点) 分布と超幾何分布との同時分布の条件つき確率 $P_r(\,|\,)$ に関する定理である：要約すれば、３種類の壷 A, B, C が $n(\geq 1)$ 個ずつあり、種類 A の壷には白玉 W が３個だけ、壷 B には W ２個と赤玉 R が１個、壷 C には W １個と R ２個が入っているが；上記 ❷ のさらなる事例として、例えば、壷を１つ無作為に選んで、玉も無作為に１個取出すと、白玉 W であったとき、その玉 W が壷 A から取出された確率計算 $P_r(A|W)$ に関する定理である。詳細は、福場 (1993, p.50) 等を参照のこと：

$$P_r(A|W) = \frac{P_r(A)\,P_r(W|A)}{\sum_{\upsilon=A,B,C} P_r(\upsilon)\,P_r(W|\upsilon)} = \frac{1/3}{1/3 + 2/9 + 1/9} = \frac{1}{2}. \qquad (1)$$

■§3. 確率樹形図について：–

これは、通常、事象の発生とともに左から右へ描かれる：まず、場合の数だけ樹の第一枝とそこに発生する確率を書き、次の第二枝には、条件つき確率を書込んでいくが；右へ進むと積の法則、上下に足すと和の法則を成立させる。例えば、前ページ ■§2. の事例では、第一枝として３本 $P_r(A) = P_r(B) = P_r(C) = 1/3$ の確率、第二枝として第一枝から２本ずつ計６本 $P_r(W|A) = 1$, $P_r(R|A) = 0$, $P_r(W|B) = 2/3$, $P_r(R|B) = 1/3$, $P_r(W|C) = 1/3$, $P_r(R|C) = 2/3$ の条件つき確率、そして右端に、積の法則による確率を以下のように描く。

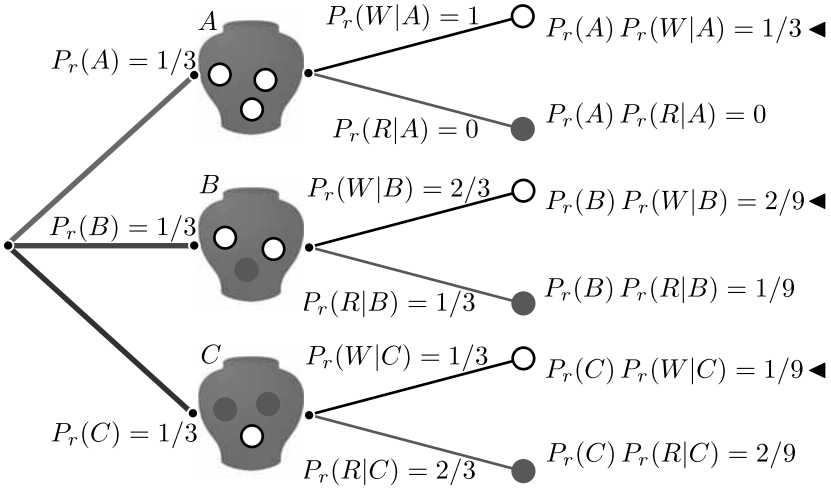


図１. 確率樹形図とベイズの定理

左図の確率樹形図から、■§2. の例題：壷を１つ無作為に選んで、玉も無作為に１個取出すと、白玉 W であったとき、その玉 W が壷 A から取出された条件つき確率 $P_r(A|W)$ の計算は、◀ 印に着目すれば、$P_r(A|W) = \frac{P_r(A \cap W)}{P_r(W)} = \frac{P_r(A)\,P_r(W|A)}{\sum_{v=A,B,C} P_r(v)\,P_r(W|v)} = \frac{1/3}{1/3+2/9+1/9} = \frac{1}{2}$ となり、視覚化して、(1) 式の計算過程を確認できる。

以上、ベイズの定理を用い、時間を遡及する (■§1.❷) には、第二枝までの試行結果が全部出尽くす必要がある。では、■§1.❶ との相違点を考えよう。例えば、『２頭の兄弟ヤギ G_{ij} がいる。そのうち１頭が黒色と判明したとき、もう１頭も黒である確率』は、右図のように、兄 i 弟 j、白 0 黒 1 とおくと、1/3 となる。よって、違いは、第二枝の有無より、❶ での事象 $A|B$ の逆 $B|A$ は意味が無い。

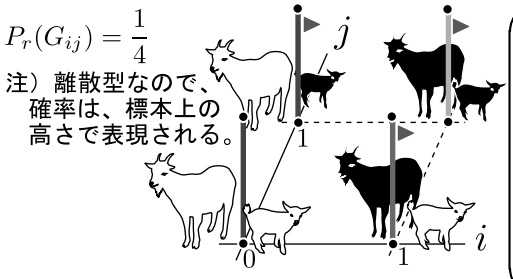


図２. 当該条件つき確率

ヤギ G_{ij} のうち１頭が黒色 B で、もう１頭も黒色の確率 $P_r(G_{11}|B)$ は、左図 ▶ 旗印に着目すれば、$P_r(G_{11}|B) = \frac{P_r(G_{11} \cap B)}{P_r(B)} = \frac{P_r(G_{11})}{\sum_{i,j \neq 0} P_r(G_{ij})} = \frac{1/4}{1/4+1/4+1/4} = \frac{1}{3}$ となる。

■§4. ３囚人問題について：–

これは、幽閉されている３囚人 A, B, C に関する問題である：要約すると、３囚人のうち、１人が (送迎車で) 釈放、残り２人は処刑されることになっているが；処刑前日、囚人 B が処刑されるとの真の情報を入手した囚人 A が、自分の釈放される主観的確率 θ を 1/3 から 1/2 へと更新した根拠は何かという問題である。詳細は、市川 (1998, p.21)、福場 (1993, p.55)、松原 (2010, p.67) 等を参照のこと。

■§5. モンティホール問題について：–

これは、ホール氏が司会を務めたテレビ番組の視聴者参加型ゲームのカーテン A, B, C に関する問題である：要約すると、３カーテンのうち、例えば１つが高級車、残り２つはヤギなどハズレ、ゲームの参加者は２回カーテンを選べることになっているが；まず１回目に、参加者がカーテン A を選んだと仮定する。次の２回目直前に、司会者からカーテン B がハズレとの情報を得た参加者が、カーテン A が当たる主観的確率 θ を 1/3 から 1/2 にしか更新しなかったために、２回目の選択もカーテン A のままとした (２回目の選択をカーテン C に変更しなかった) 根拠は何かという問題である。詳細は、市川 (1998, p.25)、今野ら (2010, p.98)、ホフマン (2000, p.252) 等を参照のこと。

■§6. ３点分布について：−

ところで、「成功: ■§4.(車で) 釈放; ■§5.(車の) 当り」の可能性は、A, B, C の３通りだから、確率変数 $\mathcal{A}_i, \mathcal{B}_i$ を考え、その「成功」の確率を θ, ρ とおくと、C のそれは $1-\theta-\rho$ となる。そこで、実現値を a_i と b_i、「成功」に 1 の値、「失敗: ■§4. 処刑; ■§5. ヤギ」には 0 の値を与えれば、C の「成功」は (0, 0) で、標本空間 $\Delta = \{(a_i, b_i)|(0, 0), (1, 0), (0, 1)\}$ を得て、それらの同時確率密度関数 $f_i(a_i, b_i)$ が、

$$f_i(a_i, b_i) = \begin{cases} \theta^{a_i}\rho^{b_i}(1-\theta-\rho)^{1-a_i-b_i} & \text{if } (a_i, b_i) \in \Delta, \\ 0 & \text{elsewhere} \end{cases} \tag{2}$$

と表現できるから、確率変数 $\mathcal{A}_i$ と $\mathcal{B}_i$ は、離散型の３点分布に従うことがわかる。

積率母関数 $M_1(t_a, t_b) \equiv E(e^{t_a\mathcal{A}_i+t_b\mathcal{B}_i}) = \sum_\Delta e^{t_a a_i+t_b b_i} f_i(a_i, b_i)$ を計算すると、$M_1(t_a, t_b) = \theta e^{t_a} + \rho e^{t_b} + 1 - \theta - \rho$ を得る。また、標本のサイズを n とすると、和の確率変数 $\mathcal{A} = \sum \mathcal{A}_i$ と $\mathcal{B} = \sum \mathcal{B}_i$ の積率母関数 $M_n(t_a, t_b) \equiv E(e^{t_a\mathcal{A}+t_b\mathcal{B}}) = \Pi_{i=1}^n M_1(t_a, t_b)$ は、$M_n(t_a, t_b) = (\theta e^{t_a} + \rho e^{t_b} + 1 - \theta - \rho)^n$ となる。よって、$0 \le a + b \le n$ を満たす非負の整数をそれぞれの実現値 a, b とおけば、$\mathcal{A}$ と $\mathcal{B}$ は３項分布に従い、以下のような同時確率密度関数 $f(a, b)$ をもつ：

$$f(a, b) = \begin{cases} \frac{n!}{a!\,b!\,(n-a-b)!}\theta^a\rho^b(1-\theta-\rho)^{n-a-b} & \text{if } 0 \le a+b \le n; \\ 0 & \text{elsewhere.} \end{cases} \tag{3}$$

標本サイズが $n = 1$ のとき、(3) 式は、(2) 式の３点分布であることに気づく。ただし、どちらの式でも、A の「成功: 車を得る」確率は $P_r(A) = \theta$、B のそれは $P_r(B) = \rho$、全事象 $P_r(A) + P_r(B) + P_r(C) = 1$ と余事象から、C の「成功」は $P_r(C) = 1 - \theta - \rho$ となる。それでは下図に、$n = 1$ の確率密度関数 $f = f(a, b)$ を示そう：$f(1, 0) = P_r(A)$; $f(0, 1) = P_r(B)$; $f(0, 0) = P_r(C)$ である。さて、詳細はのちほど述べるとして、ここで、問題 ■§4. と ■§5. 同様、B の「失敗; ここはヤギ」$\overline{B}$ を仮定すると、A が「成功」する確率 $P_r(A|\overline{B})$ は $\theta/(1-\rho)$ となり、同様に確からしい ($\theta = \rho = 1/3$) かぎり、その値は 1/2 である。

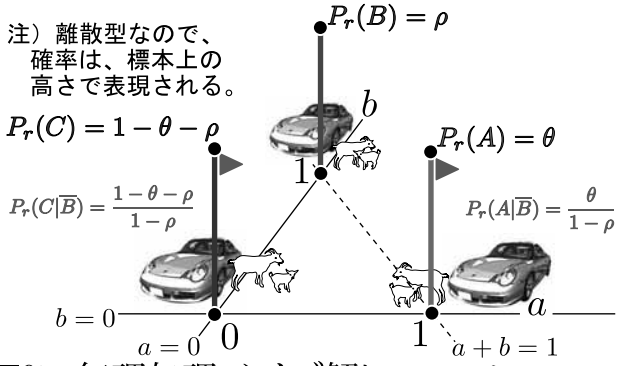
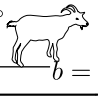


図３. 条件つき密度関数 $f(a|b)$ と $P_r(A|\overline{B})$

ヤギ ($b = 0$) が出現、B の「成功」がない $\overline{B}$ 条件下で A が「成功」する確率 $P_r(A|\overline{B})$ は、左図 ►► 印から、$P_r(A|\overline{B}) = \frac{P_r(A\cap\overline{B})}{P_r(\overline{B})} = \frac{P_r(A)}{\sum_{v\neq B} P_r(v)} = \frac{P_r(A)}{P_r(A)+P_r(C)} = \frac{1/3}{1/3+1/3} = \frac{1}{2}$ となる。



$b = 0$

■§7. 無理矢理ベイズ解について：−

図３から、■§4. ■§5. の諸問題も図２同様、ベイズの定理は使えない。あえて、右の樹形図を描くならば、B にヤギ G_B が出現後、A の「成功」する確率 $P_r(A|G_B)$ は 1/3 となる。

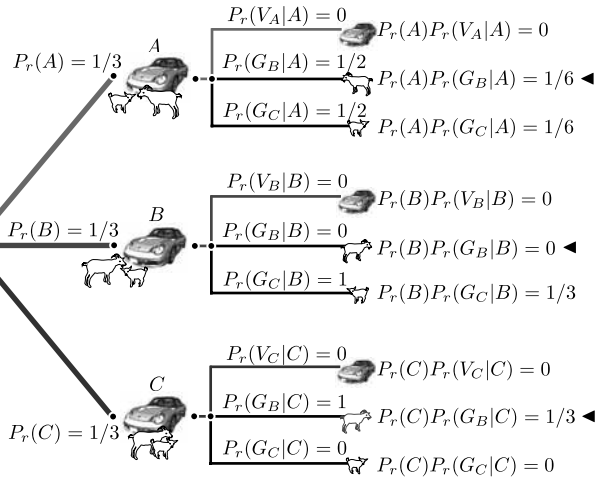


図４. 無理矢理ベイズ解

右図 ◄ 印に着目、ベイズの定理を用いると、$P_r(A|G_B) = \frac{P_r(A\cap G_B)}{P_r(G_B)} = \frac{P_r(A)\,P_r(G_B|A)}{\sum_{v=A,B,C} P_r(v)\,P_r(G_B|v)} = \frac{1/6}{1/6+0+1/3} = \frac{1}{3}$ となるが、第二枝分布、遡及の仕方、◄ 印以外の積の値等矛盾点が多い。

図４の根拠には選択数がある (今野ら, 2010, p.98)：３選択肢から A を選んだままステイするよりも、例えば、B のヤギ以外の A, C ２つから C へ選択を変える方がアタリやすいと考える。なぜならば、図４では、$P_r(V_A|A) = P_r(V_B|B) = P_r(V_C|C) = 0$ 等矛盾はさておき、$P_r(C|G_B) = \frac{1/3}{1/6+0+1/3} = \frac{2}{3} > P_r(A|G_B) = \frac{1}{3}$ を得るからである。無論、選択肢が増え、A, B, C, D, E, F の６つから A を選んだら、B, D, E, F がヤギで、２回目に A, C が残る場合は、C がアタリやすいに決まっている。しかし、問題 ■§4. ■§5. の本質はそうではない。増えた６つから、例えば、A, B を１回目に選び、D, F がヤギならば、２回目は、残る A, B, C, E からアタリを２つ (例：車体；鍵) 獲得するなど、当否の割合を同じにする必要がある。すると、ヤギの D, F がアタリの在り処を教えるはずもなく、A, B から C, E へ選択を総入換えしても、アタリやすくなっていないことに気づく。結局、選択の数や「失敗」の情報をもとに、確率樹形図を無理やり描いても、意味がない。

■§8. 新ベイズ解について：−

図３において懸案の条件つき確率 $P_r(A|\overline{B}) = \theta/(1-\rho)$ の新ベイズ解は、(3) 式のパラメータ θ, ρ を、それぞれ、ディリクレ分布〜その確率密度関数として、

$$\pi(\theta,\, \rho) = \begin{cases} \frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)}\theta^{\alpha-1}\rho^{\beta-1}(1-\theta-\rho)^{\gamma-1} & \text{if } 0 \le \theta \le 1-\rho \le 1; \\ 0 & \text{elsewhere} \end{cases} \tag{4}$$

を持つ確率変数 $\mathcal{Q}, \mathcal{R}$ の実現値とすると、$E\{\mathcal{Q}/(1-\mathcal{R})|\mathcal{A}, \mathcal{B}\} = \frac{a+\alpha}{n-b+\alpha+\gamma}$ となり、

$$E\{\frac{\mathcal{Q}}{1-\mathcal{R}}|\mathcal{A}, \mathcal{B}\} = \frac{n-b}{n-b+\alpha+\gamma}\,\frac{a}{n-b} + \frac{\alpha+\gamma}{n-b+\alpha+\gamma}\,\frac{\alpha}{\alpha+\gamma} \tag{5}$$

のような最尤推定量 $a/(n-b)$ とベータ分布の平均 $\alpha/(\alpha+\gamma)$ との凸結合である。

紙面の関係上、簡単に導出を示す。まず、(3) 式の積率母関数 $M_n(0,\, t_b)$ から、周辺分布の密度関数 $f_b = \frac{n!}{b!(n-b)!}\rho^b(1-\rho)^{n-b}$ を得て、条件つきのそれ $f(a|b)$ は、

$$f(a|b) = \frac{f(a,\, b)}{f_b} = \begin{cases} \frac{(n-b)!}{a!\,(n-a-b)!}(\frac{\theta}{1-\rho})^a(\frac{1-\theta-\rho}{1-\rho})^{n-a-b} & \text{if } a = 0,\, 1,\, \cdots,\, n-b; \\ 0 & \text{elsewhere} \end{cases} \tag{6}$$

のように、二項分布と関係し、当該 $\frac{\theta}{1-\rho}$ の最尤推定量は $\frac{a}{a+n-a-b}$ となる。そして、(2) 式とネイマンの因数定理より、核 κ_1 が $\kappa_1(a,\, b|\,\theta, \rho) \propto f_1 f_2 \cdots f_n$ となり、確率変数 $\mathcal{A}, \mathcal{B}$ は、θ, ρ の結合十分統計量なので、(4) 式の事後分布の確率密度関数 $\pi(\theta, \rho|a,\, b) \propto \kappa_1(a,\, b|\,\theta, \rho)\pi(\theta,\, \rho)$ は、$\pi(\theta, \rho|a,\, b) = \frac{\Gamma(n+\alpha+\beta+\gamma)}{\Gamma(a+\alpha)\Gamma(b+\beta)\Gamma(n-a-b+\gamma)}$ $\theta^{a+\alpha-1}\ \rho^{b+\beta-1}\ (1-\theta-\rho)^{n-a-b+\gamma-1}$ である。その変換 $\zeta = \frac{\theta}{1-\rho}, \eta = 1-\rho$ から、

$$\pi(\zeta|a,\, b) = \begin{cases} \frac{\Gamma(n-b+\alpha+\gamma)}{\Gamma(a+\alpha)\Gamma(n-a-b+\gamma)}\zeta^{a+\alpha-1}(1-\zeta)^{n-a-b+\gamma-1} & \text{if } 0 \le \zeta \le 1; \\ 0 & \text{elsewhere} \end{cases} \tag{7}$$

となり、(5) 式の期待値と事前 $(n = a = b = 0)$ ベータ分布の平均 $\frac{\alpha}{\alpha+\gamma}$ とを得る。

参考文献

市川伸一. 1998. 『確率の理解を探る：３囚人問題とその周辺』, 共立出版, 21-112.
今野紀雄, 友野典男, 田村秀. 2010. 『Newton 別冊「偶然」にひそむ数学法則、確率に強くなる：４まちがいやすい確率・統計』, ニュートンプレス, 90-113.
福場庸. 1993. 『意思決定論の基礎』, 現代数学社, 50-3, 55-8.
松原望. 2010. 『よくわかる最新ベイズ統計の基本と仕組み』, 秀和システム, 66-9.
ホフマン, ポール. 2000. 平石律子訳『放浪の天才数学者エルデシュ』, 草思社, 252-67.