Bayesian inference with ecological applications

第一章　ベイズ推定へのいざない

農学研究院　山浦悠一

１．頻度論者とベイジアンの違い
頻度論者とベイジアンの間では、確率 p の不確実性の表現が異なる。
・頻度論者：p は定数である。p の 95%信頼区間は[0.621,0.968]である。
・ベイジアン：p は一定の確率分布に従う確率変数である。[0.660 < p < 0.959]である確率は 95%である。

２．ベイズの定理は、確率 p の情報を更新する枠組みを提供する。

1.1 イントロ

私たちは長年にわたって多くの生態学者と仕事をしてきた。その結果がこの本である。

これまでの統計トレーニングは「古典的」もしくは「頻度パラダイム」とよばれ、推定、仮説検証、線形モデル、モデル選択、実験デザインに重きを置いてきた。

果たしてベイズ統計は取り組む価値があるのか？深解な沼のようなものではないのか？一つの答えは、「ベイズアプローチは有用だから魅力的」というもの。シンプルゆえに複雑なモデルに取り組むことができる。

頻度アプローチとベイズアプローチの比較をしてみよう。

1.2 Thomas Bayes

Thomas Bayes は 1701 年にロンドンに生まれ、1719 年にエディンバラ大学に入学した。論理学と神学を修めた。1742 年に王立協会のフェローに選出される。

1.2.1 チャンスの教義

ベイズの仕事の中で大きなものは、死後出版（1763 年）されたチャンスの教義に関するエッセイである。

ここでは「未知のイベント」が発生する確率について述べられている。未知のイベントは、現在のベルヌーイ試行である。「成功」もしくは「失敗」、「表」か「裏」か、「死亡する」か「生存する」か、という二値のランダムなイベントである。

「単一の試行でイベントが発生する確率」はベルヌーイ試行の成功パラメーターと呼ばれ、p で記述される。もしコインが「公正」ならば、「表」が出る確率は $p = 1/2$ である。重要なのは、p の詳細である。

われわれはしばしば p を知ることはできない。そんなときは、データに p を語らせたいところである。ベイズのエッセイは、データから次のような結論を導き出すための手法を探した：「p が 0.23 ～ 0.59 の間にある確率は 0.80 である」。確率の両端を a と b とするならば、問題は、p が a と b の区間内に存在する確率「$\Pr(a \le p \le b)$」を決定することである。

まるでベイズは p の信頼区間を描こうとしているかのようだ。いや疑いもなく、ベイズ自身の目標は、信頼区間を計算する頻度論者と同一である。未知の量の不確実性を定量化することである。しかし、ベイズの不確実性の考慮の仕方は頻度論者とは大きく異なる。

Fig. 1.1 のような、変形した 50 セントの硬貨がある。20 回トスしたところ、17 回表が出た。このデータのもでは、$p > 0.50$ だと信じたくなる。おそらく 0.70 よりも大きいだろう。ベイズのエッセイで書かれた手法は、p の不確実性を次のように定量化する。

$0.660 \le p \le 0.959$ だというチャンスは 95%

多くの人はこの主張を無理なく受け入れるだろう。しかし頻度論者は、p は一定 fixed 量だとし、p に関する確率について語ることは意味をなさない。p は(0.660, 0.959)の区間にあるかどうかである。$0.660 < p < 0.959$ という主張に確率を付与することは、円周率 π に関して $\Pr(3.0 < \pi < 3.2) = 0.95$ というくらい意味がないのである。未知の確率 p は、一定量なのである。変形したコインの一定の特徴で、ランダムではありえない。確率がある区間内に存在する確率を述べることは意味をなさない。

1.3 ベイジアンと頻度論者の違い – 「確率」の哲学

頻度論者は、確率を仮想的な繰り返されるデータセットの要約としてしか用いない。ベイジアンは、確率をすべての未知量を記述するために用いる。ただし、どちらのパラダイムが正しいとか誤っているという問題ではなく、この違いは単なる定義の問題である。

頻度論者はベイズの記述方法に満足しないかもしれないが、頻度論者の手法はベイズアプローチとよく似ている。私たちのデータに基づけば、頻度論者の p の「正確な 95%信頼区間」は[0.621, 0.968]である。これの意味するところは次のようなものであ

る：「もし多くの統計家が変形したコインを 20 回トスして 95%信頼区間をそれぞれが計算すれば、少なくとも 95%の信頼区間は p の真の値を含むだろう」。

ベイズアプローチの区間は[0.660, 0.959]だということを思い出そう。今回のような単純な問題ならば、ベイジアンと頻度論者は類似の結論に達する。ならば、なぜこの違いに思い悩まされなければならないのか？その理由は次のようである。

1. 単純さ。ベイズアプローチは複雑なデータおよびモデルに対して分かりやすい様式で取り組むことができる。時に頻度論者的手法では取り組めないことがある。それは階層モデルに特にあてはまる。

2. 正確さ。ベイズアプローチは、数学的に無理のない解析手法を漸近的な近似を必要とせずに提供する。そしてサンプルサイズが小さく限られたデータを扱う際に、正確な推定システムを提供する。

3. 一貫性。ベイズ推定は自己矛盾がなくシンプルで無理がない。パラメーターや潜在変数や予測といったすべての未知量は確率変数として扱われる。未知量の既存の知識は確率分布を用いて数学的に明示される。そのため、ベイズ推定は事前知識を更新する正式なメカニズムを提供する。そして不確実性のすべての源を適切に考慮できる。

1.4 Thomas Bayes の推定システムの基礎

ベイズ推定の基本的なツールは彼の名を冠した定理にある。ベイズの定理は確率に関する単純な記述である。ベイズ推定の勉強は条件付き確率とベイズの定理から始まる。

1.4.1 条件付き確率

Table 1.1 は、1987/88 ~ 1995/96 年の間にニュージーランド沖で操業した 238 のイカ・エビ漁船のデータである。希少で絶滅が危惧されるニュージーランドアシカ 45 個体が誤捕獲され、41 個体が死亡した。3 カ国の漁船合計 238 体からランダムに船をサンプルしよう。ロシア船が選ばれる確率は 0.517 = 123/238 である。公正な賭けのオッズは 123（ロシア船の数）:115（日本とニュージーランド船の数）、つまり 1.07:1 となる。もしあなたがロシア船に賭けた場合、あなたは私に対して$1.07 賭けることになる。

その際に、私たちはこの賭けは誤捕獲をした船を当てる賭けだと知らされたとしよう。各国の誤捕獲の確率も知ってしまったとする：日本は 5%(1/19)、ニュージーラ

ンドは 6%(6/96)、ロシアは 19%(23/123)。ベイズの定理は、このような追加的な情報の扱い方を記述する。

TABLE 1.1 Cross-classification of vessels fishing off the coast of New Zealand, 1987/88 through 1995/96 seasons, by country of origin and accidental sea lion bycatch.

	Japan	NZ	Russia	Total
No Bycatch	18	90	100	208
Bycatch	1	6	23	30
Total	19	96	123	238
Pr(Nation)	0.080	0.403	0.517	
Pr(Bycatch|Nation)	0.053	0.063	0.187	
Pr(Nation|Bycatch)	0.033	0.200	0.767	
Odds Ratio	0.397	0.370	3.072	

各国の誤捕獲率は Pr(Bycatch|Nation)として計算されている。ここで Pr(A|B)とは、B という条件の下での A が生じる確率である。これらは条件付き確率と呼ばれ、一定の制約の下での確率計算ということになる。例えば Pr(Bycatch|Russian) = 0.187 は 123 のロシア漁船のみに注目しており、残りの 115 船は全く考慮していない。

サンプルされた船が誤捕獲する確率から掛け金を決めるべきだろう。誤捕獲していた船がロシア船である確率は Pr(Russian|Bycatch) = 0.767 (23/30)である。オッズは 23:7 つまり 3.29:1 となり、あなたは私たちに対して$3.29 賭けていることになる。ここで重要なのは、Pr(Bycatch|Russian)は Pr(Russian|Bycatch)は全く異なる量だということである。

Bycatch という追加情報は、ロシア船だという情報を 0.517 から 0.767 に、オッズを 123/115 から 23/7 に変えた。この変化はオッズ比という形で要約される：(23/7) / (123/115) = 3.07。オッズ比は追加情報を得た際の掛け金の増加を示す因数である。

1.4.2 ベイズの定理

同時確率 joint probability とは、イベントのリストが生じる確率である。イベント B と A の同時確率は Pr(B, A)で表現する。

同時確率 Pr(B, A)と条件付き確率 Pr(B|A)の関係はベイズの定理の基礎である。Table 1.1 を例に挙げると、ランダムに抽出した船がロシア船でかつ誤捕獲をしている確率は 23/238 で、誤捕獲している船のうちロシア船である確率は 23/30 である。これらの比は次のようで、

$$\frac{\Pr(\text{Russian}, \text{Bycatch})}{\Pr(\text{Russian}|\text{Bycatch})} = \frac{\frac{23}{238}}{\frac{23}{30}} = \frac{30}{238}$$

ランダムに抽出した船が誤捕獲をした確率である。これを整理しなおすと、

$$\Pr(\text{Russian}, \text{Bycatch}) = \Pr(\text{Russian}|\text{Bycatch}) \Pr(\text{Bycatch})$$

となる。これは、次の基本的な関係を意味する。

$$\Pr(\mathrm{B}, \mathrm{A}) = \Pr(\mathrm{B}|\mathrm{A}) \Pr(\mathrm{A}) \tag{1.1}$$

ここで A と B を入れ替えると次のようになる。

$$\Pr(\mathrm{A}, \mathrm{B}) = \Pr(\mathrm{A}|\mathrm{B}) \Pr(\mathrm{B}) \tag{1.2}$$

Pr(B, A)と P(A, B)は同じ情報であるから、式 1.1 の右辺を式 1.2 の左辺で置き換えると、次が導かれる。

$$\Pr(\mathrm{B}|\mathrm{A}) \Pr(\mathrm{A}) = \Pr(\mathrm{A}|\mathrm{B}) \Pr(\mathrm{B})$$

そして両辺を Pr(A)で割ると、次を得る。

$$\Pr(\mathrm{B}|\mathrm{A}) = \frac{\Pr(\mathrm{A}|\mathrm{B}) \Pr(\mathrm{B})}{\Pr(\mathrm{A})} \tag{1.3}$$

ここでイベント B を複数のイベントに分割しよう：$\mathrm{B_j}$, $\mathrm{j} = 1, 2, \ldots, \mathrm{k}$。j のうち一つしか発生しないが、必ず一つは発生する。先ほどの誤捕獲の例では、$\mathrm{k} = 3$ で $\mathrm{B_j} =$ 日本、ニュージーランド、ロシアとなる。この場合、イベント $\mathrm{B_j}$ は互いに排他的で必ず発生し、確率の基本概念と式 1.2 に従い、次にようになる。

$$\Pr(\mathrm{A}) = \sum_{\mathrm{j}} \Pr(\mathrm{A}, \mathrm{B_j}) = \sum_{\mathrm{j}} \Pr(\mathrm{A}|\mathrm{B_j}) \Pr(\mathrm{B_j}) \tag{1.4}$$

式 1.3 の B に特定のイベント $\mathrm{B_j}$ を代入し、式 1.3 の右辺の分母に式 1.4 を代入することによってベイズの定理を得ることができる。

$$\Pr(B_i|A) = \frac{\Pr(A|B_i)\Pr(B_i)}{\sum_j \Pr(A|B_j)\Pr(B_j)} \tag{1.5}$$

推定の正式なメカニズムとしてのベイズの定理

ベイズの定理は自然の特定の状態の確率 B_i を A の観察の視点から計算するレシピを提供する。例えば次のような単純な積を計算することができる。

$$\Pr(A|B_i)\Pr(B_i)$$

この積をすべての状態の積の和で除す（つまり式 1.5 の右辺の分母）。これにより、式 1.5 は次のように表現される。

$$\Pr(B_i|A) \propto \Pr(A|B_i)\Pr(B_i)$$

ここでシンボル∝は「以下に比例する」という意味である。

先程の変形したコインの例に戻ろう。$p = \Pr(\text{Heads})$が 0.55 もしくは 0.90 であることを知っているとする。この成功率が製造過程で保証された 2 つのコインを持っている。しかしコインの一つは保証ラベルがない。どちらの成功率が保障されているかを決定するために、コインを 20 回トスし、表が 17 回出たとする。つまり $A = 17$ である。

標準的な確率計算を用いると（2 章で述べる）、2 つの条件に対応する確率は以下のようになる：$\Pr(A|p = 0.55) = 0.0040$、そして $\Pr(A|p = 0.90) = 0.1901$。次の 2 例は、どちらの状態かを推定するためにベイズの定理を応用したものである。

シナリオ 1

コインを受け取ったとき、受け取ったコインがどちらかに関する不確実性は $\Pr(p = 0.55) = 1/2$ と表現できるかもしれない。最初に受け取ったコインが公正なコインのトスと等しいかどうかを判断する過程をモデル化する。まず次の 2 つの値を計算する。

$$\Pr(A|p = 0.55)\Pr(p = 0.55) = 0.0040 \times \frac{1}{2} = 0.0020$$

$$\Pr(A|p = 0.90)\Pr(p = 0.90) = 0.1901 \times \frac{1}{2} = 0.0951$$

そして式 1.5 を用い、次のように結論付ける。

$$\Pr(p = 0.55|A) = \frac{0.0020}{0.0020 + 0.0951} = 0.0206$$

ちなみに Pr(p=0.90|A)=0.9794 で、Pr(p=0.55|A)との和は 1 となる。したがって、20 回トスをして 17 回表が出たという条件の下では、p=0.90 のオッズは 47：1 よりも大きくなる（下参照）。

参考：
p=0.90 のオッズ 0.9794/(1-0.9794)=47.543
p=0.55 のオッズ 0.0206/(1-0.0206)=0.02103

シナリオ 2

コインの製造会社に問い合わせれば、次のような返答が返ってくるだろう：「p = 0.55 のコインを送ったことに 95%の自信があります」。この返答を Pr(p=0.55)=0.95 と解釈すると、

$$\Pr(p = 0.55|A) = \frac{0.0040(0.95)}{0.1901(0.05) + 0.0040(0.95)} = 0.2859$$

となる。この場合、こう返答するだろう：「そうはおっしゃいますが、私たちの観察に基づくと、p = 0.90 のコインのオッズは 2.5：1 よりも大きいです。」

参考：
Pr(p = 0.55|A)のオッズは 0.2859/(1-0.2859)=0.40036 である。1:0.40036 = 2.497752:1。ちなみに Pr(p=0.90|A) = 0.71439 で、オッズは 0.71439/(1-71439) = 2.5013 となる。

アザラシとコインの例はベイズの定理が推定ツールとして有効だということを示している。まず自然の状態の不確実性に関して確率的な主張を行なう（ランダムに抽出された漁船の国籍、変形したコインの成功率 p = Pr(Heads)）。次に、自然の状態に基づいて確率の結果を記載する（誤捕獲率、n 回のトスのうち「表」が出た回数）。ベイズの定理は観察された結果に基づいて自然の状態の不確実性に関する主張を訂正するために用いられる。

本章では、ベイズ解析に関する簡単な例題を示した。ベイズ解析の理解をさらに深めるために、確率分布と尤度概念を 2 章で整理する。