【土木学会論文集　第362号／I-4（討議・回答）1985年10月】

長　尚著 "ベイズの定理の適用について" への討議

（土木学会論文集　第350号／I-2　1984年10月掲載）

▶討議者（*Discussion*）――藤野陽三（東京大学）
By Yozo FUJINO

次の2点について疑問をもちましたので討議とさせていただきます．なお，ベイズの定理は確率論の一定理であり，主観確率を導入したいわゆるベイズ統計学はベイズ定理を礎石としますが，両者は同一の概念ではないと考えます．著者はベイズの定理＝ベイズ統計学として使われているようですが，以下の討議では区別して使うことにします．

（1）「ベイズ統計学は余り有効ではない」というのが本論文の主要な結論になっていると理解されます．著者は2，3の例を通じてこのことを示そうとされているわけですが，杭の問題（p.305），つぼの問題（p.306）（確率分布のパラメーター推定の問題（p.308）は若干異なりますが，この問題については（2）参照），いずれも事前分布として散漫分布（diffused distribution），すなわち，推定したい変数に対して事前に何の統計的情報を持ちあわせていない状況を設定しています．

ベイズ統計学における事後確率は，特に比較的少ない統計データしか得られない場合，事前確率に依存するところが大きいわけです．事前確率が適切（確かなもの）であれば事後確率はより改善されますが，事前情報として何ら持ちあわせていない，あるいは事前確率が的外れであればなかなか真の値は収束しない，つまり，ベイズ統計学の有効性はケースバイケースであるということは広く認識されていることです．

ベイズ統計学の有効性が発揮されにくい状況を設定した例題からその非有効性を指摘したところで普遍的事実を導いたことにはなり得ず，かえってベイズ統計学に対する誤解を読者に与えることになると考えます．「ベイズ統計学には一般に少ない情報から実り多い結果が引き出されることが期待されている」と著者は述べておられますが，「期待し得る場合がある」と理解するのが正しいのではないでしょうか．

ベイズ統計学では，経験，周辺知識に基づいて，ある事象に対する情報を主観確率という形でとり込み，それを統計的事実（データ）とシステマティックに組み合わせて情報を更新できるのに対し，古典統計学では統計的事実のみを対象とし，主観的情報を持ちあわせていても何らそれをとり込むすべがありません．この点ベイズ統計学の方が柔軟であると考えます．特に，土木工学のように豊富な経験に基づいた技術者の勘を発揮し得る部分が少なからず存在する分野においては，ベイズ統計学の有効性は高く，文献9），10）のような研究成果は支持されるべきものと思われます．さきに述べたようにベイズ統計学の有効性は事前分布に左右されるわけで，有効と考えられる場合に適宜使っていくという姿勢が好ましいのではないでしょうか．たしかに事前分布を設定するのはやさしいことではないとは思いますが，何人かのエクスパートによる答えを集約して事前分布を作成するとか，幾通りかの事前分布を想定してその結果に対する敏感性を調べるとか，ベイズの統計学を使えるものにしていく手段はいろいろ考えられるものと思っております．

（2）細かい問題ですが，3．の「確率分布のパラメーター推定への適用について」で示されている例題（p.308）の設定そのものに疑問をもちます．

事前情報として n_1 個の観測値が与えられている場合を想定しているようですが，このときにはもはや主観確率の入り込む余地のない，すなわちベイズ統計学が適用し得ない問題となっているのではないでしょうか．

a）仮に分散 σ^2 が既知である正規母集団分布の平均値 μ の事前分布 $N(\mu', \sigma'^2)$ で与えられているとしてベイズ流に考えてみます．$\mu'=\overline{X}_1$, $\sigma'=\sigma/\sqrt{n_1}$ を用いれば，n_2 個の観測値が得られた段階での平均値のベイズ推定値 μ_B はベイズの定理を用いて

$$\mu_B=\frac{n_1\overline{X}_1+n_2\overline{X}_2}{n_1+n_2} \qquad \cdots\cdots(29')$$

となり，著者のいうところの古典的推定法による平均値 μ_C と一致します．なお，正規母集団分布の分散 σ^2 が既知でない場合も平均値，分散の事前分布として共役形を

仮定すれば式（29'）を得ることができます．とすると，論文中の式（29）は一体何を意味することになるのかが疑問として生じますが，分散 σ_1^2，σ_2^2 の扱い方に問題があると思われます．さて，式（29'）の n_1，x_1 は主観確率に関するパラメーターですから客観情報ではないわけで，古典的推定法による μ_C は式（30）ではなく

$$\mu_C=\overline{X}_2 \cdots\cdots (30')$$

とするのが正しいかと思われます．

b）あるいは次のように考えることができるかもしれません．もともとは μ に関して何の情報もない状態（diffused prior）で n_1 個の観測値が得られたとすれば平均値のベイズ推定値は当然 $\mu_B=\overline{X}_1$，そこへ n_2 個のデータが追加されたとすれば推定値はさらに更新されて $\mu_B=(n_1\overline{X}_1+n_2\overline{X}_2)/(n_1+n_2)$ となることが正規分布の場合容易に示すことができ，これは式（29'）と一致します．このときの古典的推定法による平均値 μ_C は μ_B と一致することは明らかです．

どのような問題設定のもとに μ_B(式(29))，μ_C(式(30))が求められたのでしょうか．

著者の問題設定そのものが討議者にとって不明確ですので困るのですが，仮にa）のように考えると，x_1，n_1 の値の与え方によりベイズ統計学の有効性がケースバイケースとなることがはっきりとしますし，b）のように考えればベイズ統計学はある条件下では古典統計学を包括することが明らかにされることになります．いずれにせよベイズ統計学が有効でないという結論を導き得る論理展開はこの例題においてもなされていないと考えられます．

（1985.4.17・受付）

▶回答者（*Closure*）

長　尚（信州大学）
By Takashi CHOU

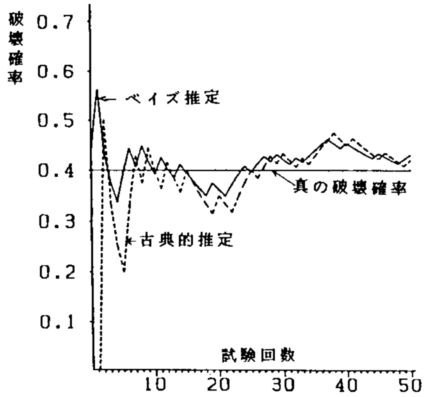
著者の論文に関し貴重なご討議を賜り心から感謝致します．ご指摘の2点について回答する前に，“著者はベイズ定理＝ベイズ統計学として”いるとありますが，拙文のまえがきの部分で「このベイズの定理には…いわゆるBayesianの立場がある．そして…異なるものである」と断ってあることを指摘しておきます．

（1）について

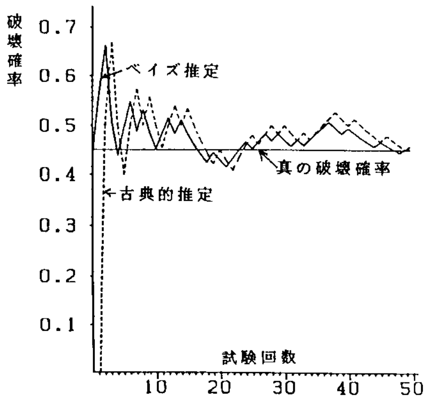
まず事前分布が散漫事前であるという“ベイズ統計学の有効性が発揮されにくい状況を設定した例題からその非有効性を指摘したところで普遍的事実を導いたことにはなり得”ない，というご指摘に対してお答えします．著者が取り上げた例題はパラメーター推定の問題（これについては（2）で触れる）の1つを除き，他の5例はすべて通常の教科書・参考書に出ている例題で，著者が有効性が発揮されにくい例題を作ったのでもなければ，ことさらにそのような問題だけをピックアップしたわけではない．しかも事前分布が散漫分布の例題だけ取り上げて議論してはいない（事前分布が散漫分布の例題はこの5例中2例である）．さらに単にそれぞれの例の非有効性を指摘したばかりではなく，なぜそうなるかについての普遍的な根拠についても考察を加えている．したがって著者は一般論として非有効性を指摘しているのである．ついでにいうと，論文で取り上げた例以外にも多数の例題について著者としては検討を加えているが，この種の問題で有効な適用例と保証できるものはみつかっていない．なお著者が“「少ない情報から実り多い結果が引き出されることが期待されている」と”記した（正確にいうとこのような表現ではなく，文脈としては，ベイズの定理を適用したものには期待されている，としているのであるが）ことに対して“「期待し得る場合がある」と理解す”べきとありますが，著者もそのように理解した上で議論している．ただし，教科書・参考書に出ている例題は，特に断りがない限り，ベイズの定理の適用の有効性が期待されているから取り上げられていると著者は考えている．この意味で事前分布が散漫分布の例題も議論の対象としているのである．特につぼの問題は「ポリヤのつぼ」といわれ，大抵の教科書・参考書でベイズの定理の説明に用いられている．そのような例が，実はその適用の効用の少ないことを承知の上で，単にこの定理の構造的内容を示すためだけにわざわざ挙げられているとは考えにくい．もしそうならなぜ適用の効用が顕著な例を示して理解を助けないのか疑問をもつ．

次に討議者は“事前確率が適切（確かなもの）であれば事後確率はより改善され”，“特に，土木工学のように豊富な経験に基づいた技術者の勘を発揮し得る部分が少なからず存在する分野においては，ベイズ統計学の有効性は高い”とされていますが，このような評価について著者は疑問を述べているのである．この点をさらに補足するために，少し長くなりますが，具体的な例題についてのシミュレーション結果を示して考察を加える．

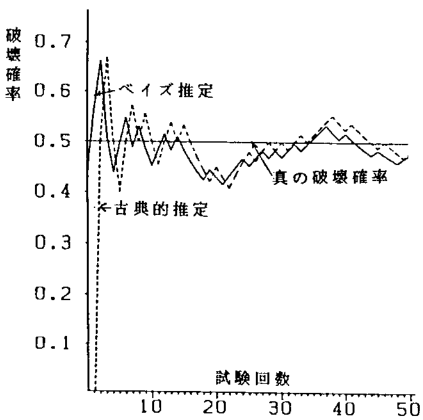
まず論文でも取り上げた，Angらの例題8.1について真の破壊確率が0.4，0.45，0.5の3ケースのシミュレーション結果を図—1〜3に示す．ただし真の破壊確率が任意の値にとれるように，事前離散確率の間隔を0.2ではなく，0.01とした．これらの図から次のことがいえる．①試験回数が3〜6回以上になると，ベイズ推定が古典的推定に勝るとは必ずしもいえない．ただし，これらの例では推定値のばらつきはベイズ推定の方がや



図―1　シミュレーションの例 1-1

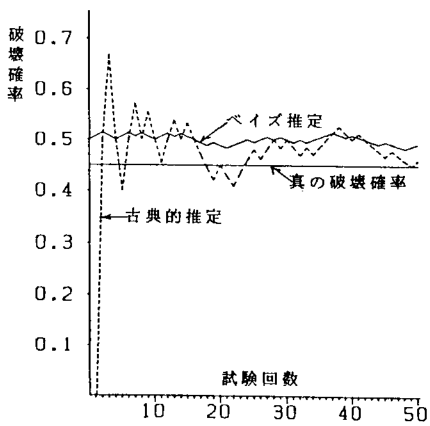


図―2　シミュレーションの例 1-2



図―3　シミュレーションの例 1-3

や少なくなっている．②事前確率が比較的適切（確かなもので，真の破壊確率が 0.40～0.50 に対して事前確率が 0.44 で，その確信レベルが 0.3～0.7（55 %））であっても，事後確率はより改善されるとは限らない．これらの例では数回の試験結果のみで古典的推定を行うことはあり得ない．したがってこのような場合ベイズの定理の有効性が発揮されるというのであろうか．しかしもし事前に推定された破壊確率がある場合には，ベイズの定理を用いるまでもなく，総合的に判断して推定するのが常識的である．討議者は“古典的統計学では……，主観的情報を持ちあわせていても何らそれをとり込むすべがありません”とあります．確かに「学」としてはとり込めませんが，技術者はそれを判断材料として利用するのが自然です．たとえば図―2 の例でいうと，1 回目の載荷試験で破壊しないということからだけでは何とも判断できませんが，事前の確率を特に変更して推定値とすることはまずないでしょう．ところがベイズ推定では，0.56 となっている．また 6 回の載荷試験で 3 回破壊したことから，通常の技術者は，破壊確率の値は 0.45～0.5 で，ほぼ推定値に近いと判断するでしょう．ところがベイズ推定値は 0.55 となっている．つまり，ベイズの定理を適用しなくても常識的な判断が可能で，場合によってはその方がベイズ推定より優れていることもあるのである．ところでこれらの例では確信レベルが低く，散漫事前分布に近い例であるから一般的でないという反論があるかもしれないので，もう 1 つこれに関連した例を挙げる．すなわち，図―4 は，事前分布が正規分布で平均値が 0.5，その確信レベルが ±0.1（90 %），真の破壊確率が 0.45 のケースのシミュレーション結果である．この例のように事前情報がかなり適切で，真の破壊確率が 90 % の確信の範囲に入っていても，データ数が数個以上になると，必ずしもベイズ推定が有効とはいえず，むしろ古典的推定の方がよいこともある．しかもこの場合は，図―1～3 の例では長所であった，ベイズ推定のばらつきの減少という性質が，かえって真の破壊確率への接近を妨げている．もちろんこの場合平均値が 0.45，その確信レベルが ±0.1（90 %），真の破壊確率が 0.45 のケースではそのシミュレーション結果を示すまでもなく，ベイズ推定は古典的推定より勝る．しかしそんなに



図―4　シミュレーションの例 1-4

精度よく事前に推定することが可能な場合はまずないし，たとえあったとしても，少ないデータによる改善の効果はなく，やはりベイズの定理の適用の効果はないことになる．一般的には図—4のようなこともあり得るし，それがどんな場合にそうなるか，推測できない状況にある．したがって討議者のいう“ベイズ統計学の有効性はケースバイケースである”としても，そのケースの判断が困難である．もし図—4の例は“事前確率が的外れで”有効性が発揮されないものであるから，例として妥当でないとされるとしたら，どのような事前確率が的外れでないといえるのか著者には理解できない．いずれにしてもこれらの例のように事前分布が散漫分布でなく，しかも事前の推定がかなり適切な場合であっても，ベイズの定理の適用の有効性は認められないのである．

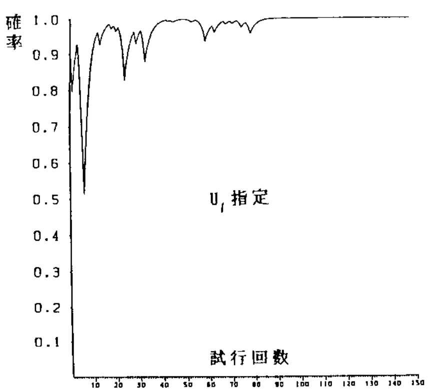
次に，つぼの問題で事前分布が散漫分布でない，シミュレーションの例を図—5～8に示す．図—5,6はU_1のつぼが指定され，事前確率を$p'_1=0.9$，$p'_2=0.05$，$p'_3=0.05$とした場合，図—7,8はU_2のつぼが指定され，事前確率を$p'_1=0.05$，$p'_2=0.9$，$p'_3=0.05$とした場合のものである．このように事前確率がかなり的を射ていても，有効性に欠ける例が図—5,7のようにあり得るのである．

最後にパラメーターの推定の問題のシミュレーション結果について示す．図—9は母集団が正規分布で真の平均値が10.0，標準偏差が3.0，事前分布が正規分布で事前の平均値の推定値とその確信レベルが，11.0±1（90％）(a) および10.5±1（90％）(b) とし，1回につき5個のデータが追加されるものとしたときの平均値の推定の例を示したものである．これらの例はいずれも，図—1～4で吟味したことと同じことがいえ，やはりベイズの定理の有効性は認められない．

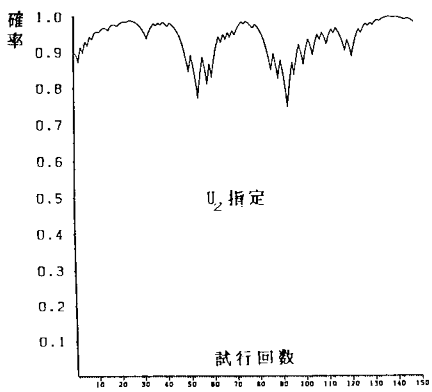
以上のシミュレーション結果に関する考察から次のことがいえる．

① 事前確率等の推定が適切でも，その確信の度合が低い場合には，少ないデータの影響を受け，必ずしもベイズ推定で改善されるとは限らない．

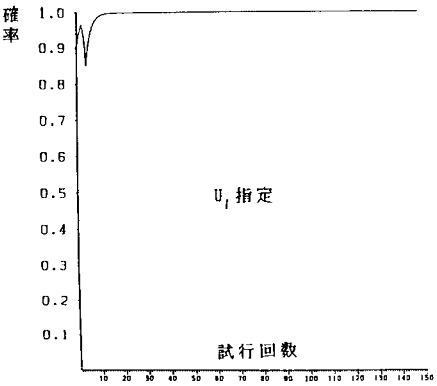
② 事前確率等の推定が適切で，その確信の度合が非常に高い場合には，事後のベイズ推定は事前推定とほとんど変わらない．このような場合は元々ベイズの定理の適用によるまでもない．



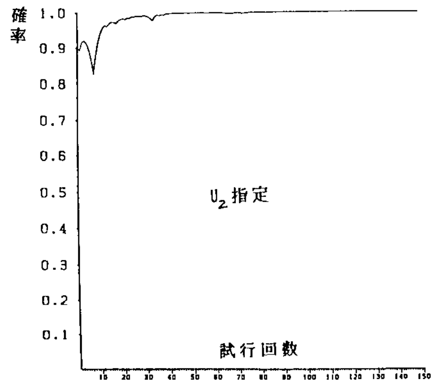
図—5 シミュレーションの例2-1



図—7 シミュレーションの例2-3



図—6 シミュレーションの例2-2



図—8 シミュレーションの例2-4

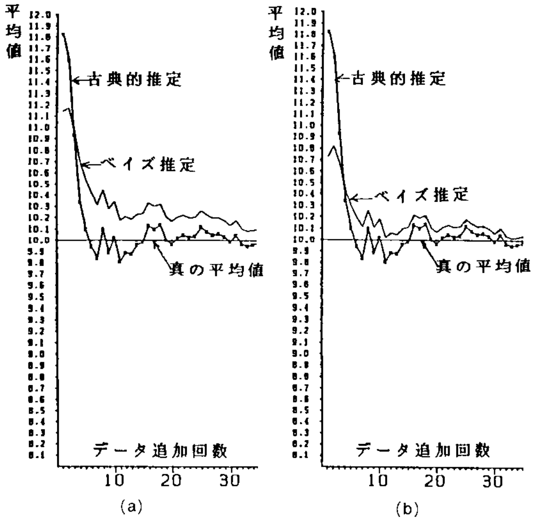


図-9 シミュレーションの例3

③ 確かに一般的にいって，ベイズ推定の方が古典的推定に比べて，推定値のばらつきは小さくなる．この傾向は事前の確信の度合の大きいほど強くなるが，一方では事前の推定のずれの影響がいつまでも残り，むしろ古典的推定の方がよくなることもある．しかもこのようなずれがあるのが普通である．

④ 主観的情報は，ベイズ統計学でなければとり込めないというような性質のものではない．

⑤ 事前確率等の推定がかなり適切でその確信の度合が高い場合でも，なかなか収束しないものもあり得る．

⑥ ある程度事前の推定が適切でその確信の度合も高い場合，すなわち討議者のいう“有効と考えられる場合”でも，上記のことからベイズ推定の方がよいというケースはかなり限られるし，そのよさの程度もそんなに顕著ではない．その上このようなケースをあらかじめ一般的に判断できるような基準はない．“豊富な経験に基づいた技術者の勘”という抽象的な保証でこのような判断が可能になるとは，ここで示した例からみても，とても考えられない．したがって有効だと思って適用しても，結果として古典的推定に劣るかもしれないし，よくてもその程度がそれほどでもないことがかなりあり得るので，ベイズの定理を適用した研究成果が有用であると保証することは一般に困難である．結論として，“ベイズ統計学の有効性は高い”という評価には大いに疑問をもつ．

なお，本論文で取り上げた例題の著者および討議者だけの考え方に異論をいっているわけではない．むしろこれらの著者・討議者の考え方の方が一般的であり，そうしたベイズ統計学への一般的評価に対して著者は疑問を提起しているのである．例題もそのような観点から典型的なものを取り上げたつもりである．

（2）について

この点に関しては討議者のご指摘どおり例題として不適切でした．“細かい問題”とはいえ，深くお詫びしますとともにご指摘に心から感謝致します．ただしパラメーター推定へのベイズの定理の適用が必ずしも有効でないという結論には，前項の回答で示した図—9の例からも明らかなように，変更はありません．まことに恐れ入りますが，本論文のこの部分は，前項の図—9とその説明文のように訂正させて頂きます．

（1985.7.11・受付）