標本調査に関するノート

標本調査に関するノート

平　館　道　子

I. 有限母集団の推測問題に対するベイジアンの観点からの一つの接近法に関して，その基本的な考え方と定式化については別掲ノートにおいて述べた*。要約すれば，大きさ N の要素からなる有限母集団において，N 次元の変量 $\boldsymbol{Y}=(Y_1,\ \cdots\cdots,\ Y_N)$ を考え，このうちある調査設計に基づいて標本 $\boldsymbol{Y}_S=(Y_1,\cdots\cdots,\ Y_{S_n})$ をとり調査した場合．これらの値が $\boldsymbol{y}_S=(y_{S_1},\ \cdots\cdots,\ y_{S_n})$ であり，測定誤差の様な非標本誤差を含まないとすれば，母集団の調査されない部分 $\boldsymbol{Y}_{\bar{S}}=(Y_{\bar{S}_1},\ \cdots\cdots,\ Y_{\bar{S}_{N-n}})$ の事後分布は，調査設計に依らず，$\boldsymbol{Y}_S=\boldsymbol{y}_S$ となるすべての $\boldsymbol{Y}$ に対して

$$f\ (\boldsymbol{y}_{\bar{S}}\mid\boldsymbol{y}_S)=\frac{f(\boldsymbol{y})}{f(\boldsymbol{y}_S)} \tag{1}$$

によって求められる。ここに $f(\boldsymbol{y})$ は $\boldsymbol{Y}$ の事前分布を示し，$f\ (\boldsymbol{y}_S)$ は $\boldsymbol{Y}_S$ の周辺分布を示す。(1) の事後分布を用いて，母集団の総計 $T=\sum_{i=1}^{N}Y_i$ あるいは平均 $M=\frac{T}{N}$ などに関する推測を行なうことができる。

もし $\boldsymbol{Y}$ が一つの共通な偶然機構 $f\ (\cdot\mid\theta)$ に従って相互に独立にその値を定めると判断されるならば，$f(\boldsymbol{y})$ の付与については，これを直接行なうのではなく，パラメーター θ の事前分布の付与を通じて行なうことができる。すなわち，θ の事前分布を $f(\theta)$ とすれば，

$$f(\boldsymbol{y})=\int\prod_{i=1}^{N}f(y_i\mid\theta)\ f(\theta)d\theta \tag{2}$$

であるから，これから $\boldsymbol{Y}_{\bar{S}}$ の事後分布は

* 平館道子：“有限母集団の推測について：ノート(I)”金沢大学法文学部論集経済学篇23．なお，同ノートの参照文献も参照されたい。

$$f(\boldsymbol{y}_{\bar{S}} \mid \boldsymbol{y}_S) = \frac{f(\boldsymbol{y})}{\int f(\boldsymbol{y})\, d\boldsymbol{y}_{\bar{S}}} \tag{3}$$

として求めることができる。（2）は N 次元変量 Y の要素を任意に置換しても変らない分布であり，この様な性格をもつ分布は exchangeable であるといわれる。exchangeable な事前分布を付与することは事前の判断において各要素を無作為化することであると云うこともできよう。

調査しようとする変量が近似的に正規分布に従ってその値を定めると判断される場合はしばしばある。以下では正規分布の場合に関するいくつかの結果をまとめよう。

II. 有限母集団の要素 Y は互いに独立に同一の正規分布 $N(\theta, v)$ に従っていると判断されるとしよう。この場合，まず分散 v についてはある値をとることができるが，θ は未知で，事前分布を付与しなければならないとしよう。正規分布の平均の事前分布については，知識を持たない状態を表現すると考えられる一様分布がとられることがあるが，ここでは正規分布 $N(m, v')$ が付与されると仮定する。（2）によって Y の事前分布は N 次元正規分布として求められ，平均，分散はすべての Y_i について等しく，それぞれ m, $v+v'$ であり，共分散も等しく v' である。従って任意の Y_i と Y_j とを置き換えても分布は変わらず exchangeable である。また共分散が 0 でないことが，標本を観測した場合，その観測値から，観測されていない他の変量の値に関する情報が与えられることを保証している。

サイズ n の標本の観測値 $\boldsymbol{y}_S$ が与えられる時，母集団の残りの部分 $Y_{\bar{S}}$ の事後分布は，（3）から $N-n$ 次元正規分布であり，共通の平均，分散，共分散をもち，それらはそれぞれ，

$$\frac{nv'\bar{y}_S+vm}{nv'+v}, \quad \frac{v\,[(n+1)v'+v]}{nv'+v}, \quad \frac{vv'}{nv'+v}$$

であることが導びかれる。ここに $\bar{y}_S$ は標本の平均を示す。従って母集団の総和 T および平均 M についてみれば，その事後分布はそれぞれ正規分布で，平均と分散は次の様になる。

$$E(T \mid \boldsymbol{y}_S) = \frac{n(Nv'+v)\bar{y}_S + (N-n)vm}{nv'+v},$$

$$V(T \mid \boldsymbol{y}_S)=\frac{(N-n)\ (Nv'+v)v}{nv'+v},$$

$$E(M \mid \boldsymbol{y}_S)=\frac{n\left(v'+\dfrac{v}{N}\right)\bar{y}_S+\left(1-\dfrac{n}{N}\right)vm}{nv'+v},$$

$$V(M \mid \boldsymbol{y}_S)=\frac{(N-n)v}{Nn}\cdot\frac{Nv'+v}{Nv'+\dfrac{N}{n}v}.$$

Mの事後平均は，観測されたデータ $\bar{y}_S$（これは M の伝統的な点推定である）と事前の判断 m との荷重平均になっている。T あるいは M に関する推測はこれらの事後分布を用いて様々に行なうことができる。例えば，母集団平均がその中に入る事後確率が 0.95 であるような最短区間を求めたければ，それは $E\ (M \mid \boldsymbol{y}_S)\pm 1.96\ \sqrt{V(M \mid \boldsymbol{y}_S)}$ である。

ここでは θ の事前分布を正規分布としたが，情報を含まないこと，あるいは知識のないことを示す分布としてしばしば使用される一様分布をとるとすると，$\boldsymbol{Y}$ の事前分布は真正な分布とならず，しかも確率分布として解釈することが困難な関数になってしまう。これは極めて不自然なことであり，推測という目的にとって不整合な事態というべきであろう。この事は慎重な考慮なしに一様分布やこれに類する，情報を含まない状態を示す分布を用いることに対する一つの警告ともなっており，結局ベイジアンの立場で推測問題を扱おうとすれば，事前に何の判断もないということでは済まされないことを示しているとも云えよう。

III. 次に θ と共に v も未知の場合について，θ と v の同時事前分布が次の様な正規—逆ガンマ分布でよく表現できると仮定しよう。

$$f(\theta,\ v)=\frac{1}{\sqrt{2\pi v}}e^{-\frac{n'(\theta-m)^2}{2v}}\cdot\left(\frac{a}{2}\right)^{\frac{p}{2}}\Gamma\left(\frac{p}{2}\right)^{-1}\ v^{-\left(\frac{p}{2}+1\right)}\ e^{-\frac{a}{2v}}. \tag{4}$$

この分布のパラメーターは $(m,\ n',\ p,\ a)$ である。この時 $(\boldsymbol{Y},\ \theta,\ v)$ の同時分布 $f(\boldsymbol{y},\ \theta,\ v)$ はその核が

$$v^{-\left(\frac{N+p+1}{2}+1\right)} \exp\left[-\frac{1}{2v}\left\{\sum_{i=1}^{N}(y_i-\theta)^2+n'(\theta-m)^2+a\right\}\right]$$

となる。Y の周辺事前分布を求めるために，まずこれを v で積分すると

$$f(\boldsymbol{y},\ \theta)=\int_0^\infty f(\boldsymbol{y},\ \theta,\ v)dv$$

$$\propto\left\{\Sigma(y_i-\theta)^2+n'(\theta-m)^2+a\right\}^{-\frac{N+p+1}{2}}$$

となる。これを更に θ で積分するために，上式を

$$\left\{Q+(N+n')\left(\theta-\frac{\Sigma y_i+n'm}{N+n'}\right)^2\right\}^{-\frac{N+p+1}{2}} \tag{5}$$

と書き改める。ここに Q は

$$(\boldsymbol{y}-\boldsymbol{m})'H(\boldsymbol{y}-\boldsymbol{m})+a$$

であり，$\boldsymbol{m}$ はすべての要素が m である N 次元ベクトル，H は対角要素が $\dfrac{N+n'-1}{N+n'}$，その他の要素がすべて $-\dfrac{1}{N+n'}$ である正値定符号 $N\times N$ 対称行列である。(5) を θ について積分すると Y の分布 $f(\boldsymbol{y})$ が得られ，その核は次のようになる。

$$f(\boldsymbol{y})\propto\left\{p+\ (\boldsymbol{y}-\boldsymbol{m})'H\ (\boldsymbol{y}-\boldsymbol{m})\frac{p}{a}\right\}^{-\frac{N+p}{2}}. \tag{6}$$

この分布は一般化された N 次元 t 分布であり，自由度は p，それぞれの Y の平均は m，分散共分散行列は $\dfrac{a}{p-2}H^{-1}$ である。この様な t 分布を以下では簡単のため単に t 分布とよぶことにする。Y_i の分散は $\dfrac{a}{p-2}\left(1+\dfrac{1}{n'}\right)$，$Y_i$ と Y_j の共分散は $\dfrac{a}{(p-2)n'}$ である。これは v が既知の場合と同様 exchangeable な事前分布である。

次に標本 $\boldsymbol{y}_S$ が与えられた時の $Y_{\bar{S}}$ の事後分布を求める。

▶一般に自由度 ϕ の K 次元 t 分布において K_2 個の変量 $\boldsymbol{t}_2$ の値が与えられた時の残りの変量（この $K-K_2=K_1$ 個の変量を $\boldsymbol{t}_1$ とあらわす）の条件つき分布は次の様に求められる。K 次元 t 分布において，その核が

$$f(\boldsymbol{t}) \propto \left\{\phi+(\boldsymbol{t}-\eta)' H_0(\boldsymbol{t}-\eta)\right\}^{-\frac{\phi+K}{2}} \tag{7}$$

で与えられる時，変量の順序を変えて $\boldsymbol{t}=(\boldsymbol{t}_1,\ \boldsymbol{t}_2)$，$\eta=(\eta_1,\ \eta_2)$ とし，H_0 をこれに従って並べ換え

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}$$

と分割しておく。ここに H_0 は $K\times K$ の正値定符号対称行列である。$\boldsymbol{t}_1$ の条件つき分布は

$$f(\boldsymbol{t}_1 \mid \boldsymbol{t}_2)=\frac{f(\boldsymbol{t})}{f(\boldsymbol{t}_2)}$$

で求められるが，$f(\boldsymbol{t}_2)$ は $f(\boldsymbol{t})$ を $\boldsymbol{t}_1$ で積分して次の量に比例する。

$$\left\{\phi+(\boldsymbol{t}_2-\eta_2)' G(\boldsymbol{t}_2-\eta_2)\right\}^{-\frac{\phi+K_2}{2}} \tag{8}$$

ここに G は，

$$G=H_{22}-H_{12}' H_{11}^{-1} H_{12} \tag{9}$$

であり，$K_2\times K_2$ 正値定符号対称行列である。従って求める条件つき分布については，（7）を（8）で除して，

$$\left\{1+\frac{(\boldsymbol{t}_1-\boldsymbol{M}_1)' H_{11}(\boldsymbol{t}_1-\boldsymbol{M}_1)}{\phi+(\boldsymbol{t}_2-\eta_2)' G(\boldsymbol{t}_2-\eta_2)}\right\}^{-\frac{\phi+K_2+K_1}{2}}$$

が分布の核となる。これは自由度 $\phi+K_2$ の K_1 次元 t 分布である。$\boldsymbol{M}_1$ は

$$\eta_1-H_{11}^{-1} H_{12}(\boldsymbol{t}_2-\eta_2)$$

であり，変量 $\boldsymbol{t}_1$ の平均は $\boldsymbol{M}_1$，分散共分散行列は

$$\frac{\phi+(t_2-\eta_2)'G(t_2-\eta_2)}{\phi+K_2-2}H_{11}^{-1}$$

である。◀

以上の結果を我々の問題に適用すると，$Y_{\bar{S}}$ の事後分布の核は，

$$\left\{1+\frac{(y_{\bar{S}}-M_{\bar{S}})'H_{\bar{S}\bar{S}}(y_{\bar{S}}-M_{\bar{S}})}{a+(y_S-m)'G_S((y-m)}\right\}^{-\frac{N+p}{2}}$$

となる。ここに $M_{\bar{S}}$ はすべての要素が $\dfrac{n\bar{y}_S+n'm}{n+n'}$ である $(N-n)$ 次元ベクトルで，$H_{\bar{S}\bar{S}}$ は（6）の $\boldsymbol{H}$ を分割した部分行列であり，G_S は（9）によって得られる。具体的には，次の様になる。

$$H_{\bar{S}\bar{S}}=\begin{pmatrix} \frac{N+n'-1}{N+n'} & & -\frac{1}{N+n'} \\ & \ddots & \\ -\frac{1}{N+n'} & & \frac{N+n'-1}{N+n'} \end{pmatrix}, \qquad (N-n)\times(N-n)$$

$$G_S=\begin{pmatrix} \frac{n+n'-1}{n+n'} & & -\frac{1}{n+n'} \\ & \ddots & \\ -\frac{1}{n+n'} & & \frac{n+n'-1}{n+n'} \end{pmatrix}, \qquad n\times n$$

従って，$Y_{\bar{S}}$ の事後分布は $N-n$ 次元 t 分布であり，自由度は $p+n$，平均はすべて等しく $\dfrac{n\bar{y}_S+n'm}{n+n'}$，分散共分散行列は

$$\frac{1}{p+n-2}\left\{a+(y_S-m)'G_S(y_S-m)\right\}H_{\bar{S}\bar{S}}^{-1}$$

で与えられる。$\dfrac{1}{p+n-2}\left\{a+(y_S-m)'G_S(y_S-m)\right\}$ を Q とおけば，$y_{\bar{S}}$ の分散は $\dfrac{n+n'+1}{n+n'}Q$，共分散は $\dfrac{1}{n+n'}Q$ である。

以上の結果から母集団の和 T の事後分布が得られる。t 分布は再生性をもつから，T の分布もやはり自由度 $p+n$ の1次元 t 分布であり，その平均，分散はそれぞれ，

$$E(T|\boldsymbol{y}_S) = \frac{n(N+n')\bar{y}_S + (N-n)n'm}{n+n'},$$

$$V(T|\boldsymbol{y}_S) = \frac{(N+n')\ (N-n)}{n+n'}\ Q$$

である。母集団平均 $M=\frac{\sum_{i=1}^{N} Y_i}{N}$ の事後分布は同様に自由度 $p+n$ の t 分布であり，その平均，分散はそれぞれ

$$E\ (M|\boldsymbol{y}_S) = \frac{n\left(1+\frac{n'}{N}\right)\bar{y}_S + (N-n)\frac{n'}{N}m}{n+n'},$$

$$V\ (M|\boldsymbol{y}_S) = \frac{(N-n)}{N^2}\ \frac{N+n'}{n+n'}\ Q$$

である。

$(\theta,\ v)$ の事前分布として正規—逆ガンマ分布（4）をとったが，これは正規分布において，θ，v が共に未知である場合の自然共役 (**natural conjugate**) な分布とよばれている。 **II.** で既に述べた通り， 推測に際して事前の判断を真正な分布に表現することが必要であるが，しかし，事前分布は常にすべての点において厳密に定めることが可能だという訳ではないから，この曖昧さ，あるいはゆとりを利用して，可能ならば解析的に都合のよい分布でこれを表現することを考えるとよいであろう。もとの分布が正規分布である場合，最も都合のよい事前分布は正規分布と逆ガンマ分布から構成される正規—逆ガンマ分布である。あるいは分散の逆数をパラメーターとして，正規—ガンマ分布を事前分布としてもよい。これらの事前分布は事後分布を尤度関数と同じかあるいは関連する分布に導くという性格を持っており，それらの分布は性格もよく知られ，解析的な取扱いにも都合がよい。事前の判断をこれらの分布によって表現しようという場合，要はパラメーターの付与が問題になるわけであるが，その分布の性格や，結果として得られる事後分布に与える影響をよく考慮する必要がある。

正規—逆ガンマ分布の場合， v が与えられた時の θ の分布は正規分布で，その平均は m，分散は $\frac{v}{n'}$ である。また θ の周辺分布 $f(\theta)$ は

$$f(\theta) \propto \left(p + \frac{pn'}{a}(\theta-m)^2\right)^{-\frac{1}{2}(p+1)}$$

となり，自由度 p，平均 m，分散は $p>2$ の時 $\frac{a}{n'(p-2)}$ の t 分布である。v の周辺分布 $f(v)$ は

$$f(v) \propto v^{-\left(\frac{1}{2}p+1\right)} e^{-\frac{a}{2v}}$$

で逆ガンマ分布であり，平均，分散はそれぞれ，$\frac{a}{p-2}$，$\frac{2a^2}{(p-2)^2(p-4)}$ である。従ってもし v について有限な分散を想定するならば，$p>4$ でなければならない。p は結果として生じる t 分布の自由度に関連するものであり，事前および事後の分散は p の減少関数になっている。p は4より大でなければならないが，標本数が小さい時にはその影響も大きく，p を大きくとる事はタイトな事後分布を得ることになるであろう。a も分散に関連する量であるが，これは p とは逆に分散を大きくする要素であり，p と a は θ，v それぞれの周辺分布の性格を考慮することによって同時に付与すべきものであろう。パラメーター n' は変量 Y が従っていると考えられる偶然機構 $N(\theta,\ v)$ の平均に関してどれだけ精密な判断が下せるかを表現するものであり，$Y_{\bar{S}}$ の事後の平均が $\frac{n\bar{y}_S+n'm}{n+n'}$ であることから見てもわかる通り，事前の判断に対するウエイトとなっているから，事前に標本いくつ分に見合うだけの情報を持っているかを示すものと考えられよう。従って付与に際しては，I. J. Good の imaginary sample の様な考え方に従えば目安が与えられ，付与という困難な作業の助けになるであろう。

$Y_{\bar{S}}$ の事後分布を求める際，(3) から

$$f(\boldsymbol{y}_{\bar{S}}\,|\,\boldsymbol{y}_S)=\frac{\int \prod_{i\in\bar{S}} f(y_i|\theta)\cdot\prod_{i\in S} f(y_i|\theta)\,f(\theta)d\theta}{f(\boldsymbol{y}_S)}$$

$$=\int \prod_{i\in\bar{S}} f(y_i|\theta)\quad f(\theta\,|\,\boldsymbol{y}_S)d\theta$$

とすれば，これは θ の事後分布による $\prod_{i\in\bar{S}} f(y_{\bar{S}}|\theta)$ の平均を求めていることになる。因みに **II.** の問題の場合，θ の事後平均は $\frac{nv'\bar{y}_S+vm}{nv'+v}$ であり，$Y_{\bar{S}}$ の事後平均に等しい。また θ の事後の分散は $\frac{vv'}{nv'+v}$ であり，$Y_{\bar{S}}$ の事

後共分散に等しい。同様な点を III. の場合についてみれば，θ，v の事後平均は，それぞれ，$\dfrac{n\bar{y}_S+n'm}{n+n'}$，$\dfrac{1}{n+p-2}[a+(\boldsymbol{y}_S-\boldsymbol{m})'G_S(\boldsymbol{y}_S-\boldsymbol{m})]$ である。後者を $E(v|\boldsymbol{y}_S)$ とあらわせば，θ，v の事後の分散はそれぞれ，$\dfrac{E(v|\boldsymbol{y}_S)}{n+n'}$，$\dfrac{2}{n+p-4}[E(v|\boldsymbol{y}_S)]^2$ である。この事から，$Y_{\bar{S}}$ の事後平均は θ のそれと等しく，事後の分散共分散行列は $E(v|\boldsymbol{y}_S)\ H_{\bar{S}\bar{S}}^{-1}$ となっていることがわかる。ここでも $Y_{\bar{S}}$ の事後共分散は θ の事後分散と等しい。

IV. 変量に関する推測を関係のある他の変量を用いて行なうことがある。例えば田の収量を耕地面積から推測するとか，米の消費量を世帯人員数から推測するなどである。この様な場合で，規模の経済性が一定であることが想定される時，これまでの結果が直ちに応用できる。

いま $(X_1,\ Y_1)$，……，$(X_N,\ Y_N)$ という大きさ N の母集団があり，X_1，……，X_N は既知であるとしよう。これは例えば，ある地域にある田全体について，各々その面積が知られている時，米の全収量の推測が問題となったり，同じくある N 世帯からなる地域において，各世帯の人数が知られている時米の全消費量を知りたい場合などである。

問題となっている変量 Y は，もう一つの量 X と

$$Y_i=\beta X_i+\varepsilon_i \qquad i=1,\ \cdots\cdots,\ N$$

という関係をもっていると想定されるとしよう。更に Y_i は互いに独立に $N(\beta X_i,\ v_i)$ に従っていると判断されるものとする。v_i については，一定と考えられる場合，あるいは各々異るが例えば X_i^2 に比例すると考えられる場合などがあろうが，もし後者の場合ならば，$Z_i=\dfrac{Y_i}{X_i}$ は $N(\beta,\ v)$ に従う。v は $v_i=X_i^2v$ をみたす正の定数である。Z_i は単位面積当り収量，1人当り消費量などであり，これは規模の経済性が一定で，β を平均として変動し，その変動の仕方も規模に依存しないことを意味する。これまでと同様に，標本を $\boldsymbol{Y}_S=(Y_{S_1},\ \cdots\cdots,\ Y_{S_n})$，残りの部分を $\boldsymbol{Y}_{\bar{S}}=(Y_{\bar{S}_1},\cdots\cdots,\ Y_{\bar{S}_{N-n}})$ とし，$\boldsymbol{Y}_S$ に対応する X を $\boldsymbol{x}_S=(x_{S_1},\ \cdots\cdots,\ x_{S_n})$，$\boldsymbol{Y}_{\bar{S}}$ に対応するそれを $\boldsymbol{x}_{\bar{S}}=(x_{\bar{S}_1},\ \cdots\cdots,\ x_{\bar{S}_{N-n}})$ とする。また $\boldsymbol{Y}_S$ の観測値をこれまでと同様 $\boldsymbol{y}_S$ とすると，

v の値が知られている場合 β の事前分布が $N(m,\ v')$ であれば，$Y_{\bar{S}}$ の事後分布は正規分布であり，$Y_{\bar{S}_i}$ の平均，分散はそれぞれ，$\dfrac{nv'\bar{z}_S+vm}{nv'+v}\,x_{\bar{S}_i}$，$\dfrac{\{(n+1)v'+v\}}{nv'+v}\,x_{\bar{S}_i}{}^2$ であり，$Y_{\bar{S}_i}$ と $Y_{\bar{S}_j}$ の共分散は $\dfrac{vv'}{nv'+v}\,x_{\bar{S}_i}\,x_{\bar{S}_j}$ である。この場合，Y の事前分布は **exchangeable** ではないが，Z のそれは **exchangeable** である。従って総和 T の事後分布は，

$$N\left(\frac{nv'\bar{z}_S+vm}{nv'+v}\sum_{i=1}^{N-n}x_{\bar{S}_i}+\sum_{i=1}^{n}y_{S_i},\right.$$

$$\left.\frac{\{(n+1)v'+v\}v}{nv'+v}\sum_i x_{\bar{S}_i}{}^2+2\frac{vv'}{nv'+v}\sum\sum_{i<j}x_{\bar{S}_i}\,x_{\bar{S}_j}\right)$$

である。

V. 次に v が未知の場合，**III.** と同様にして，$(\beta,\ v)$ の事前分布を（4）とすれば，事後分布に関しては，

$$\frac{(p+n)\ (\boldsymbol{y}_{\bar{S}}-\boldsymbol{M}_{\bar{S}})'\ H_{\bar{S}\bar{S}}\ (\boldsymbol{y}_S-\boldsymbol{M}_{\bar{S}})}{a+(\boldsymbol{y}_S-\boldsymbol{M}_S)'\ G(\boldsymbol{y}_S-\boldsymbol{M}_S)}$$

が自由度 $p+n$ の $N-n$ 次元 t 変量である。$\boldsymbol{M}_S$ は $m\boldsymbol{x}_S$，$\boldsymbol{M}_{\bar{S}}$ は $\dfrac{n\bar{z}_S+n'm}{n+n'}\,\boldsymbol{x}_{\bar{S}}$ であり，$H_{\bar{S}\bar{S}}$ はその $(i,\ i)$ 要素が $\dfrac{N+n'-1}{N+n'}\,x_{\bar{S}_i}{}^2$，$(i,\ j)$ 要素が $-\dfrac{1}{N+n'}\,x_{\bar{S}_i}x_{\bar{S}_j}$ の $(N-n)\times(N-n)$ 行列であり，G は $(i,\ i)$ 要素が $\dfrac{n+n'-1}{n+n'}x_{S_i}{}^2$，$(i,\ j)$ 要素が $-\dfrac{1}{n+n'}x_{S_i}\ x_{S_j}$ の $n\times n$ 行列であり，共に正値定符号である。これから総和 T は同様に t 分布に従い，事後の平均と分散はそれぞれ，

$$\sum_{i=1}^{n}y_S+\frac{n\bar{z}_S+n'm}{n+n'}\sum_{i=1}^{N-n}x_{\bar{S}_i},$$

$$\frac{1}{p+n-2}\Big[a+(\boldsymbol{y}_S-\boldsymbol{M}_S)'\,G(\boldsymbol{y}_S-\boldsymbol{M}_S)\Big]\left[\frac{(n+n'+1)\sum x_{\bar{S}_i}{}^2+2\sum\sum_{i>j}x_{\bar{S}_i}x_{\bar{S}_j}}{n+n'}\right]$$

である。

VI. v_i が一定である場合，これを v としよう。この場合には，単位規模当りの生産性は，β を平均として変動するが，その変動の仕方は規模の 2 乗に逆比例すると想定されることを意味する。この場合にはこれまでの結果を直ちに応用することはできないが，同様に分析することができる。v が知られている場合，$\boldsymbol{Y}$ の事前分布は正規分布で平均が $m\boldsymbol{x}$，分散共分散行列は，その (i, i) 要素が $x_i^2 v' + v$，(i, j) 要素が $x_i x_j v'$ である。ここに $\boldsymbol{x} = (x_1, \cdots\cdots, x_N)$ である。従ってこの分布は exchangeable ではない。$\boldsymbol{Y}_{\bar{S}}$ の事後分布は

$$N\left(\frac{v'\sum x_S y_S + vm}{v'\sum x_S^2 + v}\,\boldsymbol{x}_{\bar{S}},\ \Sigma\right)$$

である。Σ は (i, i) 要素が $\dfrac{v\ [v'(x_{\bar{S}i}{}^2 + \sum x_S^2) + v]}{v'\sum x_S^2 + v}$，$(i, j)$ 要素が $\dfrac{x_{\bar{S}i}\,x_{\bar{S}j} v v'}{v'\sum x_S^2 + v}$ の $(N-n)\times(N-n)$ 行列である。従って T は平均が

$$\sum y_S + \frac{(N-n)\ (vm + v'\sum x_S y_S)\bar{x}_{\bar{S}}}{v'\sum x_S^2 + v},$$

分散が

$$\frac{v(N-n)\ [v'\{\sum x_S^2 + (N-n)\bar{x}_{\bar{S}}{}^2\} + v]}{v'\sum x_S^2 + v}$$

の正規分布である。以上において $\bar{x}_{\bar{S}}$ は $x_{\bar{S}}$ の平均を，$\sum x_S^2$，$\sum x_S y_S$ は S についての和をあらわす。

VII. v が未知の場合，(β, v) に対して（4）と同様な事前分布をとるとすると，$\boldsymbol{y}_{\bar{S}}$ の事後分布は自由度 $p+n$ の t 分布でその平均は

$$\frac{n'm + \sum x_S y_S}{\sum x_S^2 + n'}\,\boldsymbol{x}_{\bar{S}},$$

分散共分散行列は，

$$\frac{1}{n+p-2}\left[a + (\boldsymbol{y}_S - m\boldsymbol{x}_S)'\ G_S(\boldsymbol{y}_S - m\boldsymbol{x}_S)\right] H_{\bar{S}\bar{S}}^{-1} \tag{10}$$

で与えられる。ここに G_S は $(i,\ i)$ 要素が $\dfrac{\sum x_S{}^2+n'-x_{Si}{}^2}{\sum x_S{}^2+n'}$, $(i,\ j)$ 要素が $-\dfrac{x_{S_i}\ x_{S_j}}{\sum x_S{}^2+n'}$ の $n\times n$ 行列であり, $H_{\bar{S}\bar{S}}$ は$(i,\ i)$要素が $\dfrac{\sum_{i=1}^{N} x_i{}^2+n'-x_{\bar{S}i}{}^2}{\sum_{i=1}^{N} x_i{}^2+n'}$, $(i,\ j)$ 要素が $-\dfrac{x_{\bar{S}i}\,x_{\bar{S}j}}{\sum x_i{}^2+n'}$ の$(N-n)\times(N-n)$行列である。(10)の $H_{\bar{S}\bar{S}}{}^{-1}$ の乗数を Q とすれば, $y_{\bar{S}_i}$ の分散は $\dfrac{\sum x_S{}^2+x_{\bar{S}i}{}^2+n'}{\sum x_S{}^2+n'}Q$, $y_{\bar{S}_i}$ と $y_{\bar{S}_j}$ の共分散は $\dfrac{x_{\bar{S}i}\,x_{\bar{S}j}}{\sum x_S{}^2+n'}Q$ である。以上の結果から, 総和 T の事後分布は同様に自由度 $p+n$ の t 分布で, 平均, 分散は次の様に与えられる。

$$\sum y_S+\frac{\sum x_S y_S+n'm}{\sum x_S{}^2+n'}(N-n)\bar{x}_{\bar{S}},$$

$$\frac{(N-n)\{\sum x_S{}^2+n'+(N-n)\ \bar{x}_{\bar{S}}{}^2\}}{\sum x_S{}^2+n'}Q.$$