スタイン型推定量の平均自乗誤差に関するパラドックス

難波，明生

(Citation)
国民経済雑誌,191(1):87-96

(Issue Date)
2005-01

(Resource Type)
departmental bulletin paper

(Version)
Version of Record

(JaLCDOI)
https://doi.org/10.24546/00055980

(URL)
https://hdl.handle.net/20.500.14094/00055980

スタイン型推定量の平均自乗誤差に関するパラドックス

難　波　明　生

計量経済学・統計学において最小自乗推定量は非常に重要かつ強力な推定量である。しかし，Stein (1956) および James and Stein (1961) により提案されたスタイン型推定量は，攪乱項が正規分布で分散が既知の場合に，最小自乗推定量を平均自乗誤差において優越する事が知られている。また，スタイン型推定量に含まれる分散を不偏推定量で置き換えれば，分散が未知の場合でも最小自乗推定量を優越する。本稿では，多変量 t 分布に従う攪乱項をもつ線形回帰モデルを考え，分散が既知の場合でも，分散を推定量で置き換えた推定量の方が，既知の分散を用いた推定量よりも小さな平均自乗誤差をもちうると言う，逆説的な結果が得られることを示す。

キーワード　最小自乗推定量，スタイン型推定量，平均自乗誤差，多変量 t 分布

1 はじめに

計量経済学・統計学においては，線形回帰モデルが頻繁に用いられる。線形回帰モデルに関しては，最小自乗推定量が最良線形不偏推定量 (Best Liner Unbiased Estimator, BLUE) であると言うことがよく知られている。しかしながら，攪乱項が正規分布に従い，かつその分散が既知である場合に，Stein (1956) および James and Stein (1961) は，最小自乗推定量を平均自乗誤差において優越する推定量を提案した。さらに彼らは，分散が未知の場合でも，推定量に含まれる分散を不偏推定量で置き換えることによって得られる推定量が，最小自乗推定量を優越することを示した。これらの推定量はスタイン型推定量 (Stein-rule Estimator) と呼ばれる。Stein (1956) および James and Stein (1961) を発端として，多くの推定量が提案され，その性質の分析が行われれてきた。

これらの分析の多くでは，攪乱項が正規分布に従うと仮定されている。しかし，実際のデータを見てみると，特に金融データなどで，正規分布よりも裾の厚い分布から得られたと考えられるデータが多く見受けられる。正規分布よりも裾の厚い分布としてよく知られているのが t 分布である。正規分布からの乖離による影響を調べるために，多くの研究において多変量 t 分布が用いられてきた。多変量 t を用いた研究としては，Zellner (1976), Prucha and

Kelejian (1984), Ullah and Zinde-Walsh (1984), Judge *et al.* (1985), Sutradhar and Ali (1986), Sutradhar (1988), Singh (1988, 1991), Giles (1991, 1992), Ohtani (1991, 1993), Ohtani and Giles (1993) and Ohtani and Hasegawa (1993), Namba (2001), Namba and Ohtani (2002) 等があげられる。さらに，Fourdrinier and Strawderman (1996) は，t分布よりも一般的な spherically symmetric な分布の平均を推定する問題を考え，一定の条件の下で，仮に分散が既知であっても，分散が未知の場合に対するスタイン方推定量を用いたほうが平均自乗誤差が改善される場合があるという，逆説的な結果が得られることを示している。

そこで本稿では，多変量t分布に従う攪乱項を持つ回帰モデルを考え，Fourdrinier and Strawderman (1996) と同様に，仮に既知であったとしても，分散を推定した方がスタイン型推定量の平均自乗誤差が改善される場合があることを示す。

2 モデルと推定量

次の回帰モデルを考える。

$$y = X\beta + u \tag{1}$$

yは$n\times1$の従属変数のベクトル，Xは$n\times k$の説明変数の行列，βは$k\times1$の回帰係数のベクトルである。また，uは$n\times1$の攪乱項のベクトルである。uに関しては，多変量正規分布が仮定されることが一般的であるが，本稿ではuは多変量t分布に従うと仮定する。したがって，uの確率密度関数は

$$p(u\,|\,\nu,\sigma) = \frac{\nu^{\nu/2}\Gamma((\nu+n)/2)}{\pi^{n/2}\Gamma(\nu/2)\sigma^n}\frac{1}{\{\nu+u'u/\sigma^2\}^{(n+\nu)/2}} \tag{2}$$

となる。ただし，σはスケールパラメータ，νは分布の自由度と呼ばれる。$\nu>2$の場合，$E[u]=0$ and $E[uu']=[\nu\sigma^2/(\nu-2)]I_n$となることが容易に証明できる。このことから，スケールパラメータσは正規分布の分散に相当するパラメータとなっていることがわかる。また，$\nu\to\infty$となるにしたがって，多変量t分布は多変量正規分布に近づいていくことが知られている。

Zellner (1976) によって示されているように，多変量t分布の確率密度関数は，以下のように書き表すことができる。

$$p(u\,|\,\nu,\sigma) = \int_0^\infty p_N(u\,|\,\tau)\,p_{IG}(\tau\,|\,\nu,\sigma)\,d\tau \tag{3}$$

ただし，

$$p_N(u\,|\,\tau) = (2\pi\tau^2)^{-n/2}\exp[-u'u/2\tau^2], \tag{4}$$

は正規分布の確率密度関数であり，

$$p_{IG}(\tau\,|\,\nu,\sigma) = \frac{2(\nu\sigma^2/2)^{\nu/2}}{\Gamma(\nu/2)}\tau^{-(\nu+1)}\exp[-\nu\sigma^2/2\tau^2]. \tag{5}$$

は逆ガンマ分布の確率密度関数である。

(1)におけるβの最小自乗推定量は

$$b=S^{-1}X'y \tag{6}$$

となる。ただし，$S=X'X$である。最小自乗推定量は最良線形不偏推定量 (Best Liner Unbiased Estimator, BLUE) であることが良く知られている。これに対し，Stein (1956) および James and Stein (1961) によって提案されたスタイン型推定量は次のように定義される。

$$b_{\mathrm{SR1}}=\left(1-\frac{a_1\sigma^2}{b'Sb}\right)b \tag{7}$$

ただし，a_1は$0\leqq a_1\leqq 2(k-2)$を満たす定数である。また，平均自乗誤差 (Mean Squared Error, MSE) を次のように定義する。

$$\begin{aligned}\mathrm{MSE}[\bar{\beta}]&=E[L(\bar{\beta};\beta)]\\&=E[(\bar{\beta}-\beta)'S(\bar{\beta}-\beta)]\end{aligned} \tag{8}$$

ただし，$\bar{\beta}$はβの任意の推定量であり，$L(\bar{\beta};\beta)=(\bar{\beta}-\beta)'S(\bar{\beta}-\beta)$は損失関数と呼ばれる。$k\geqq 3$かつ$u$が正規分布に従う場合 ($\nu\to\infty$ の場合)，最小自乗推定量が BLUE であるにもかかわらず，

$$\mathrm{MSE}[b_{\mathrm{SR1}}]\leqq\mathrm{MSE}[b] \tag{9}$$

が成立する。また，$a_1=(k-2)$ の時に(7)式の推定量の平均自乗誤差は最小となるので，ここからは$a_1=(k-2)$を用いることにする。

しかし，(7)式の推定量は，分散（あるいはスケールパラメータ）σが未知である場合には用いることができない。そこで，σ^2をその不偏推定量$\hat{\sigma}^2=e'e/(n-k)$で置き換えることにより，次の推定量が得られる。

$$b_{\mathrm{SR2}}=\left(1-\frac{a_2e'e}{b'Sb}\right)b \tag{10}$$

ただし，$e=y-Xb$は残差のベクトル，$0\leqq a_2\leqq 2(k-2)/(n-k+2)$である。この推定量に関しても，$k\geqq 3$かつ$u$が正規分布に従う場合，

$$\mathrm{MSE}[b_{\mathrm{SR2}}]\leqq\mathrm{MSE}[b] \tag{11}$$

という関係が成り立つ。(7)，(10)式の推定量はスタイン型推定量と呼ばれる。$a_2=(k-2)/(n-k+2)$ の時に(10)式の平均自乗誤差は最小となるので，今後$a_2=(k-2)/(n-k+2)$を用いる。また，(3)式の関係を用いれば，(9)，(11)が，攪乱項が多変量t分布の場合にも成立することを示すのは容易である。

ところで，(10)式では(7)式の未知パラメータが推定量で置き換えられているので，我々は

$$\mathrm{MSE}[b_{\mathrm{SR1}}] \leqq \mathrm{MSE}[b_{\mathrm{SR2}}] \tag{12}$$

という関係が成り立つと期待するのが自然である。しかし，本稿の目的は，ある条件の下で，

$$\mathrm{MSE}[b_{\mathrm{SR1}}] > \mathrm{MSE}[b_{\mathrm{SR2}}] \tag{13}$$

という逆説的な結果が得られるということを示すことである。次節では上のような関係が成立し得る事を解析的に証明する。

3 平均自乗誤差

b_{SR1}，b_{SR2} の平均自乗誤差は，それぞれ次のように定義される。

$$\mathrm{MSE}[b_{\mathrm{SR1}}] = E[(b_{\mathrm{SR}_1} - \beta)' S (b_{\mathrm{SR}_1} - \beta)] \tag{14}$$

$$\mathrm{MSE}[b_{\mathrm{SR2}}] = E[(b_{\mathrm{SR}_2} - \beta)' S (b_{\mathrm{SR}_2} - \beta)] \tag{15}$$

従って，

$$H(p, q) = E[(b' Sb)^p (e' e)]^q \tag{16}$$

$$J(p, q) = E[(b' Sb)^p (e' e)^q \beta' Sb] \tag{17}$$

と定義すれば，

$$\mathrm{MSE}(b_{\mathrm{SR1}}) = H(1, 0) - 2a_1\sigma^2 + a_1^2\sigma^4 H(-1, 0) - 2J(0, 0) + 2a_1\sigma^2 J(-1, 0) + \beta' S\beta \tag{18}$$

$$\mathrm{MSE}(b_{\mathrm{SR2}}) = H(1, 0) - 2a_2 H(0, 1) + a_2^2 H(-1, 2) - 2J(0, 0) + 2a_2 J(-1, 1) + \beta' S\beta \tag{19}$$

となる。また，Appendix で証明されているように，$H(p, q)$ および $J(p, q)$ は以下のようになる。

$$H(p, q) = (\sigma^2)^{p+q} \nu^{\nu/2} \sum_{i=0}^{\infty} \left(\frac{\gamma^i}{i!}\right) (\gamma + \nu)^{p+q-i-\nu/2} \times \frac{\Gamma((n-k)/2+q)\Gamma(\nu/2-p-q+i)\Gamma(k/2+p+i)}{\Gamma((n-k)/2)\Gamma(\nu/2)\Gamma(k/2+i)} \tag{20}$$

$$J(p, q) = (\sigma^2)^{p+q} \theta \nu^{\nu/2} \sum_{i=0}^{\infty} \left(\frac{\gamma^i}{i!}\right) (\gamma + \nu)^{p+q-i-\nu/2} \times \frac{\Gamma((n-k)/2+q)\Gamma(\nu/2-p-q+i)\Gamma(k/2+p+i+1)}{\Gamma((n-k)/2)\Gamma(\nu/2)\Gamma(k/2+i+1)} \tag{21}$$

ただし，$\gamma = \theta/\sigma^2$，$\theta = \beta' S\beta$ である。このことを用いれば，以下の定理が得られる。

定理 1 $\gamma=0$ かつ $3<\nu<n-k+2$ であれば，$\mathrm{MSE}[b_{\mathrm{SR1}}] > \mathrm{MSE}[b_{\mathrm{SR2}}]$ が成立する。逆に，$\gamma=0$ かつ $n=k+2 \leqq \nu$ ならば，$\mathrm{MSE}[b_{\mathrm{SR1}}] \leqq \mathrm{MSE}[b_{\mathrm{SR2}}]$ が成立する。

証明：

(20)，(21) において $\gamma=0$ とすれば

$$H(p, q) = (\sigma^2)^{p+q} \nu^{p+q} \frac{\Gamma((n-k)/2+q)\Gamma(\nu/2-p-q)\Gamma(k/2+p)}{\Gamma((n-k)/2)\Gamma(\nu/2)\Gamma(k/2)} \tag{22}$$

$$J(p, q) = 0 \tag{23}$$

が得られる。さらに，$a_1=(k-2)$ および $a_2=(k-2)/(n-k+2)$ を用いて，若干の計算を行えば，

$$\frac{\mathrm{MSE}[b_{SR1}]-\mathrm{MSE}[b_{SR2}]}{\sigma^2}=(k-2)\left[-1+\frac{\nu(n-k)}{(n-k+2)(\nu-2)}\right] \tag{24}$$

が得られる。つまり，

$$\frac{\nu(n-k)}{(n-k+2)(\nu-2)}>1 \tag{25}$$

ならば $\mathrm{MSE}[b_{SR1}]>\mathrm{MSE}[b_{SR2}]$ が成立する。$\nu-2>0$, $n-k+2>0$ に注意して(25)を解けば $\nu<n-k+2$ が得られる。同様に，$n-k+2\leqq\nu$ ならば，$\mathrm{MSE}[b_{SR1}]\leqq\mathrm{MSE}[b_{SR2}]$ となる。

(証明終)

定理1により，攪乱項の自由度 ν がある程度小さければ（$\nu<n-k+2$），$\gamma=0$ の点において $\mathrm{MSE}[b_{SR1}]>\mathrm{MSE}[b_{SR2}]$ となる事が示された。また，攪乱項が正規分布に従う時（$\nu\to\infty$ の時），$\gamma=0$ の点においても，$\mathrm{MSE}[b_{SR1}]<\mathrm{MSE}[b_{SR2}]$ となることが分かる。しかし，$\gamma\neq0$ の点での分析を解析的に行うことは非常に困難である。従って，次節では，数値計算により b_{SR1}，b_{SR1} の平均自乗誤差を比較する。

4　数値計算による分析

本節では，前節で導いた平均自乗誤差の公式を用い，数値計算により b_{SR1}, b_{SR2} の平均自乗誤差を比較する。数値計算においては，相対的な平均自乗誤差 $\mathrm{MSE}[\bar{\beta}]/\mathrm{MSE}[b]$ の値を計算する。ただし，$\bar{\beta}$ は β の任意の推定量である。相対的な平均自乗誤差の値が1よりも小さければ，$\bar{\beta}$ は最小自乗推定量よりも小さな平均自乗誤差を持つことになる。数値計算には Fortran を用い，$H(p, q)$ および $J(p, q)$ の無限級数に関しては，増分が 10^{-12} よりも小さくなった時に，収束したとみなすことにする。(20)および(21)を用いれば，$\mathrm{MSE}[b_{SR1}]/\mathrm{MSE}[b]$，$\mathrm{MSE}[b_{SR2}]/\mathrm{MSE}[b]$ の値は n, k, ν および γ に依存することが容易に証明できる。したがって，数値計算においては以下の値を用いた。

- $k=3, 4, 5, 6, 8.$
- $n=20, 30, 40.$
- $\nu=3, 5, 10, 15, 20, 30, \infty.$
- $\gamma\geqq0$ を満たすさまざまな値。

得られた結果は表1のとおりである。代表的な結果として $k=5$, $n=20$, $\nu=5, 10, 15, 20$ の場合の結果を示している。表1より b_{SR1}, b_{SR2} 共に，1より小さい相対的平均自乗誤差を持つ事がわかる。この事は，攪乱項に多変量 t 分布を仮定した場合でも，スタイン型推定量が最小自乗推定量よりも小さな平均自乗誤差を持つことを表している。また，定理1より，$n-k+$

表1 相対的平均自乗誤差の値（$k=5$, $n=20$ の場合）

γ	$\nu=5$		$\nu=10$		$\nu=15$		$\nu=20$	
	b_{SR1}	b_{SR2}	b_{SR1}	b_{SR2}	b_{SR1}	b_{SR2}	b_{SR1}	b_{SR2}
0.0	0.64000	0.47059	0.52000	0.47059	0.48000	0.47059	0.46000	0.47059
0.1	0.64429	0.47685	0.52758	0.47894	0.48889	0.47964	0.46959	0.47998
0.2	0.64852	0.48294	0.53497	0.48706	0.49754	0.48843	0.47891	0.48912
0.3	0.65269	0.48887	0.54218	0.49495	0.50597	0.49698	0.48798	0.49799
0.4	0.65679	0.49463	0.54920	0.50263	0.51417	0.50530	0.49681	0.50663
0.6	0.66480	0.50572	0.56274	0.51738	0.52992	0.52126	0.51375	0.52320
0.8	0.67253	0.51624	0.57561	0.53136	0.54488	0.53638	0.52982	0.53889
1.0	0.68000	0.52626	0.58788	0.54462	0.55908	0.55072	0.54506	0.55377
1.5	0.69751	0.54929	0.61607	0.57499	0.59159	0.58348	0.57988	0.58772
2.0	0.71347	0.56987	0.64116	0.60187	0.62033	0.61239	0.61059	0.61762
3.0	0.74125	0.60517	0.68366	0.64724	0.66863	0.66087	0.66199	0.66761
4.0	0.76444	0.63443	0.71810	0.68394	0.70735	0.69971	0.70296	0.70744
5.0	0.78400	0.65916	0.74642	0.71415	0.73886	0.73134	0.73612	0.73970
6.0	0.80066	0.68039	0.77000	0.73939	0.76486	0.75748	0.76332	0.76620
7.0	0.81500	0.69885	0.78988	0.76076	0.78657	0.77936	0.78593	0.78824
8.0	0.82746	0.71507	0.80682	0.77905	0.80492	0.79789	0.80493	0.80680
9.0	0.83837	0.72946	0.82140	0.79487	0.82057	0.81374	0.82108	0.82259
10.0	0.84800	0.74232	0.83405	0.80866	0.83405	0.82744	0.83492	0.83615
12.0	0.86422	0.76437	0.85487	0.83153	0.85602	0.84982	0.85734	0.85817
14.0	0.87734	0.78265	0.87124	0.84968	0.87307	0.86729	0.87463	0.87520
16.0	0.88816	0.79807	0.88441	0.86440	0.88664	0.88124	0.88831	0.88870
18.0	0.89724	0.81128	0.89521	0.87656	0.89767	0.89263	0.89938	0.89965
20.0	0.90496	0.82274	0.90422	0.88677	0.90680	0.90207	0.90850	0.90868
22.0	0.91160	0.83279	0.91183	0.89545	0.91446	0.91002	0.91613	0.91625
24.0	0.91738	0.84168	0.91834	0.90292	0.92098	0.91680	0.92260	0.92268
26.0	0.92246	0.84961	0.92398	0.90942	0.92659	0.92265	0.92816	0.92820
28.0	0.92694	0.85673	0.92890	0.91511	0.93146	0.92774	0.93298	0.93300
30.0	0.93094	0.86316	0.93323	0.92013	0.93574	0.93221	0.93720	0.93720
40.0	0.94578	0.88787	0.94886	0.93845	0.95106	0.94827	0.95227	0.95223
50.0	0.95537	0.90469	0.95860	0.94997	0.96050	0.95821	0.96152	0.96147
100.0	0.97633	0.94457	0.97883	0.97423	0.97992	0.97871	0.98047	0.98044
150.0	0.98389	0.96050	0.98579	0.98266	0.98654	0.98572	0.98692	0.98689

2=17であるので，$\gamma=0$ の点で，$n=5, 10, 15$ の時には MSE$[b_{SR2}]<$MSE $[b_{SR1}]$ となり，$n=20$ の時には MSE$[b_{SR2}]>$MSE$[b_{SR1}]$ となるはずである。表1からも，この結果が成り立つことが確認できる。さらに，表1はこの結果が $\gamma\neq0$ の場合にも成り立つことを示している。また，ν が小さいほど，MSE$[b_{SR2}]$ の平均自乗誤差は，MSE$[b_{SR1}]$ の平均自乗誤差と比較して小さくなる。

さらに，$(k, n, \nu)=(5, 20, 17)$ の場合には，$\gamma=0$ の点で，MSE$[b_{SR2}]$=MSE$[b_{SR2}]$ となるはずである。同様に，$\gamma=0$ の点で，MSE$[b_{SR1}]$=MSE$[b_{SR2}]$ となる場合についての数値計算を行った。$(k, n, \nu)=(3, 20, 19)$, $(4, 20, 18)$, $(5, 20, 17)$, $(8, 20, 14)$ の場合の結果は

表２　相対的平均自乗誤差の値（$n=20$ の場合）

γ	$k=3,\ \nu=19$		$k=4,\ \nu=18$		$k=5,\ \nu=17$		$k=8,\ \nu=14$	
	b_{SR1}	b_{SR2}	b_{SR1}	b_{SR2}	b_{SR1}	b_{SR2}	b_{SR1}	b_{SR2}
0.0	0.70175	0.70175	0.55556	0.55556	0.47059	0.47059	0.35714	0.35714
0.1	0.71048	0.71047	0.56527	0.56527	0.47980	0.47980	0.36397	0.36396
0.2	0.71887	0.71885	0.57468	0.57467	0.48877	0.48875	0.37066	0.37065
0.3	0.72695	0.72691	0.58380	0.58376	0.49749	0.49746	0.37723	0.37720
0.4	0.73472	0.73466	0.59263	0.59257	0.50598	0.50592	0.38368	0.38363
0.6	0.74940	0.74928	0.60948	0.60935	0.52230	0.52217	0.39622	0.39612
0.8	0.76302	0.76283	0.62531	0.62510	0.53777	0.53757	0.40832	0.40815
1.0	0.77567	0.77539	0.64020	0.63990	0.55245	0.55216	0.41998	0.41973
1.5	0.80352	0.80301	0.67374	0.67317	0.58603	0.58548	0.44740	0.44691
2.0	0.82683	0.82611	0.70272	0.70188	0.61569	0.61485	0.47255	0.47178
3.0	0.86305	0.86197	0.74987	0.74854	0.66543	0.66405	0.51702	0.51565
4.0	0.88926	0.88794	0.78612	0.78441	0.70520	0.70336	0.55504	0.55309
5.0	0.90861	0.90717	0.81449	0.81253	0.73748	0.73530	0.58784	0.58536
6.0	0.92318	0.92171	0.83706	0.83495	0.76405	0.76161	0.61639	0.61345
7.0	0.93437	0.93291	0.85530	0.85311	0.78619	0.78358	0.64142	0.63810
8.0	0.94310	0.94170	0.87024	0.86803	0.80484	0.80213	0.66353	0.65988
9.0	0.95003	0.94869	0.88264	0.88045	0.82073	0.81796	0.68317	0.67926
10.0	0.95561	0.95435	0.89306	0.89090	0.83439	0.83159	0.70074	0.69661
12.0	0.96394	0.96284	0.90946	0.90743	0.85658	0.85382	0.73077	0.72635
14.0	0.96978	0.96883	0.92171	0.91982	0.87375	0.87108	0.75548	0.75089
16.0	0.97404	0.97322	0.93115	0.92940	0.88739	0.88483	0.77614	0.77147
18.0	0.97728	0.97656	0.93861	0.93700	0.89844	0.89601	0.79365	0.78895
20.0	0.97980	0.97917	0.94464	0.94316	0.90757	0.90526	0.80867	0.80399
22.0	0.98182	0.98126	0.94962	0.94825	0.91522	0.91304	0.82168	0.81705
24.0	0.98348	0.98297	0.95378	0.95251	0.92172	0.91965	0.83306	0.82850
26.0	0.98485	0.98440	0.95731	0.95613	0.92731	0.92535	0.84309	0.83862
28.0	0.98602	0.98560	0.96035	0.95924	0.93216	0.93030	0.85200	0.84762
30.0	0.98701	0.98663	0.96298	0.96194	0.93641	0.93464	0.85997	0.85568
40.0	0.99042	0.99015	0.97223	0.97144	0.95162	0.95020	0.88972	0.88592
50.0	0.99240	0.99219	0.97778	0.97714	0.96097	0.95980	0.90910	0.90572
100.0	0.99626	0.99616	0.98889	0.98857	0.98018	0.97956	0.95169	0.94960
150.0	0.99752	0.99746	0.99259	0.99238	0.98672	0.98630	0.96712	0.96563

表２のとおりである。表２より，γが大きくなるにしたがって，b_{SR2} の平均自乗誤差が b_{SR1} の平均自乗誤差よりも小さくなることがわかる。

以上の結果より，$\nu<n-k+2$ であれば，σが既知であったとしても，σ^2 の代わりに不偏推定量 $\hat{\sigma}^2$ を用いた推定量 b_{SR2} の方が，b_{SR1} よりも小さな平均自乗誤差を持つという，ある種逆説的な結果が得られる。

Appendix

ここでは$H(p, q)$, $J(p, q)$ の公式(20)と(21)を導く。まず，$E[(b'Sb)^p(e'e)^q|\tau]$ の公式を導出する。

$u_1=b'Sb/\tau^2$, $u_2=e'e/\tau^2$ とすると，τが所与の時，$u_1\sim\chi'^2_k(\lambda)$, $u_2\sim\chi^2_{n-k}$ となる。ただし，$\chi'^2_k(\lambda)$ は自由度k，非心度λの非心カイ２乗分布を表し，$\lambda=\beta'S\beta/\tau^2$ である。さらに，u_1とu_2は独立である。従って

$$E[(b'Sb)^p(e'e)^q\,|\,\tau]=(\tau^2)^{p+q}\sum_{i=0}^{\infty}K_i\int_0^{\infty}\int_0^{\infty}u_1^{k/2+p+i-1}u_2^{(n-k)/2+q-1}\exp\left[-\frac{u_1+u_2}{2}\right]du_1du_2 \tag{26}$$

となる。ただし，

$$K_i=\frac{w_i(\lambda)}{2^{n/2+i}\Gamma((n-k)/2)\Gamma(k/2+i)} \tag{27}$$

$w_i(\lambda)=\exp(=\lambda/2)(\lambda/2)^i/i!$ である。$\nu_1=u_1/u_2$, $\nu_2=u_2$ という変数変換を行うと，(26)式の積分は

$$\int_0^{\infty}\int_0^{\infty}\nu_1^{k/2+p+i-1}\nu_2^{n/2+p+q+i-1}\exp\left[-\frac{(1+\nu_1)\nu_2}{2}\right]d\nu_1d\nu_2 \tag{28}$$

となる。さらに，$z=\dfrac{(1+\nu_1)\nu_2}{2}$ と言う変数変換を行えば，(28)式は

$$2^{n/2+p+q+i}\Gamma(n/2+p+q+i)\int_0^{\infty}\frac{\nu_1^{k/2+P+i-1}}{(1+\nu_1)^{n/2+p+q+i}}d\nu_1 \tag{29}$$

となる。さらに，$t=\nu_1/(1+\nu_1)$ という変数変換を行えば，

$$E[(b'Sb)^p(e'e)^q\,|\,\tau]=(2\tau^2)^{p+q}\sum_{i=0}^{\infty}w_i(\lambda)\frac{\Gamma(k/2+p+i)\Gamma((n-k)/2+q)}{\Gamma((n-k)/2)\Gamma(k/2+i)} \tag{30}$$

が得られる。

つぎに，$E[(b'Sb)^p(e'e)^q\beta'Sb|\tau]$ に対する公式を導出する。(30)式をβで微分すると

$$\begin{aligned}
&\frac{\partial E[(b'Sb)^p(e'e)^q\,|\,\tau]}{\partial\beta}\\
&=(2\tau^2)^{p+q}\sum_{i=0}^{\infty}\frac{w_i(\lambda)}{\partial\beta}\frac{\Gamma(k/2+p+i)\Gamma((n-k)/2+q)}{\Gamma((n-k)/2)\Gamma(k/2+i)}\\
&=(2\tau^2)^{p+q}\sum_{i=0}^{\infty}\left[-\frac{S\beta}{\tau^2}w_i(\lambda)+\frac{S\beta}{\tau^2}w_{i-1}(\lambda)\right]\frac{\Gamma(k/2+p+i)\Gamma((n-k)/2+q)}{\Gamma((n-k)/2+i)}\\
&=-\frac{S\beta}{\tau^2}E[(b'Sb)^p(e'e)^q\,|\,\tau]\\
&\quad+\frac{S\beta}{\tau^2}(2\tau^2)^{p+q}\sum_{i=0}^{\infty}w_i(\lambda)\frac{\Gamma(k/2+p+i+1)\Gamma((n-k)/2+q)}{\Gamma((n-k)/2)\Gamma(k/2+i+1)}
\end{aligned} \tag{31}$$

となる。また，bおよび$e'e$を用いれば$E[(b'Sb)^p(e'e)^q|\tau]$ は次のように書き表すことができる。

$$E[(b'Sb)^p(e'e)^q\,|\,\tau]=\int_0^{\infty}\int_{-\infty}^{\infty}(b'Sb)^p(e'e)^q p_N(b\,|\,\tau)p_e(e'e\,|\tau)\,db\,d(e'e) \tag{32}$$

ただし，$p_e(e'e|\tau)$ はτを所与としたときの$e'e$の確率密度関数であり，

$$p_N(b\mid\tau)=\frac{1}{(2\pi)^{k/2}\mid\tau^2S^{-1}\mid^{1/2}}\exp\left[-\frac{(b-\beta)'S(b-\beta)}{2\tau^2}\right] \tag{33}$$

である。(33)を用いれば,

$$E[(b'Sb)^p(e'e)^q\mid\tau]\partial\beta=\frac{1}{\tau^2}E[(b'Sb)^p(e'e)^qSb\mid\tau]-\frac{S\beta}{\tau^2}E[(b'Sb)^p(e'e)^q\mid\tau] \tag{34}$$

が得られる。(31)式と(34)式を比較することにより,

$$E[(b'Sb)^p(e'e)^q\beta'Sb\mid\tau]=\beta'S\beta(2\tau^2)^{p+q}\sum_{i=0}^{\infty}w_i(\lambda)\frac{\Gamma(k/2+p+i+1)\Gamma((n-k)/2+q)}{\Gamma((n-k)/2)\Gamma(k/2+i+1)} \tag{35}$$

が得られる。

(3)式より,

$$H(p,q)=\int_0^\infty E[(b'Sb)^p(e'e)^q\mid\tau]p_{IG}(\tau\mid\nu,\sigma)d\tau \tag{36}$$

$$J(p,q)=\int_0^\infty E[(b'Sb)^p(e'e)^q\beta'Sb\mid\tau]p_{IG}(\tau\mid\nu,\sigma)d\tau \tag{37}$$

であり，$\nu_3=(\theta+\nu\sigma^2)/2\tau^2$ という変数変換を行えば，(20)および(21)が得られる。

参考文献

Fourdrinier, D., and W. E. Strawderman, 1996, "A Paradox Concerning Shrinkage Estimators: Should a Known Scale Parameter Be Replaced by an Estimated Value in the Shrinkage Factor?" *Journal of Multivariate Analysis*, 59, 109-140.

Giles, J. A., 1991, "Pre-testing for Linear Restrictions in a Regression Model with Spherically Symmetric Disturbances," *Journal of Econometrics*, 50, 377-398.

Giles, J. A., 1992, "Estimation of the Error Variance after a Preliminary-test of Homogeneity in a Regression Model with Spherically Symmetric Disturbances," *Journal of Econometrics*, 53, 345-361.

James, W., and C. Stein, 1961, "Estimation with Quadratic Loss," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, 361-379. University of California Press, Berkeley.

Judge, G., S. Miyazaki, and T. Yancey, 1985, "Minimax Estimators for the Location Vectors of Spherically Symmetric Densities," *Econometric Theory*, 1, 409-417.

Namba, A., 2001, "MSE Performance of the 2SHI Estimator in a Regression Model with Multivariate *t* Error Terms," *Statistical Papers*, 42, 81-96.

Namba, A., and K. Ohtani, 2002, "MSE Performance of the Double *k*-class Estimator of Each Individual Regression Coefficient under Multivariate *t*-Errors," in Aman Ullah, Alan T. K. Wan, and Anoop Chaturvedi (ed.), *Handbook of Applied Econometrics and Statistical Inference*. chap. 15, 305-326, Marcel Dekker.

Ohtani, K., 1991, "Small Sample Properties of the Interval Constrained Least Squares Estimator When Error Terms Have a Multivariate *t* Distribution," *Journal of the Japan Sta-*

tistical Society, 21, 197-204.

Ohtani, K., 1993, "Testing for Equality of Error Variances between Two Linear Regressions with Independent Multivariate t Errors," *Journal of Quantitative Economics*, 9, 85-97.

Ohtani, K., and J. Giles, 1993, "Testing Linear Restrictions on Coefficients in a Linear Regression Model with Proxy Variables and Spherically Symmetric Disturbances," *Journal of Econometrics*, 57, 393-406.

Ohtani, K., and H. Hasegawa, 1993, "On Small Sample Properties of R^2 *in a Linear Regression Model with Multivariate* t Errors and Proxy Variables," *Econometric Theory*, 9, 504-515.

Prucha, I. R., and H. H. Kelejian, 1984, "The Structure of Simultaneous Equation Estimators: A Generalization towards Nonnormal Disturbances," *Econometrica*, 52, 721-736.

Singh, R. S., 1988, "Estimation of Error Variance in Linear Regression Models with Error Having Multivariate Student-t Distribution with Unknown Degrees of Freedom," *Economics Letters*, 27, 47-53.

Singh, R. S., 1991, "James-Stein Rule Estimators in Linear Regression Models with Multivariate-t Distributed Error," *The Australian Journal of Statistics*, 33, 145-158.

Stein, C., 1956, "Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution," in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, 197-206. University of California Press, Berkeley.

Sutradhar, B. C., 1988, "Testing Linear Hypothesis with t Error Variable," *Sankhyā, Series B, Indian Journal of Statistics*, 50, 175-180.

Sutradhar, B. C., and M. M. Ali, 1986, "Estimation of the Parameters of a Regression Model with a Multivariate t Error Variable," *Communications in Statistics-Theory and Methods*, 15, 429-450.

Ullah, A., and V. Zinde-Walsh, 1984, "On the Robustness of LM, LR, and W Tests in Regression Models," *Econometrica*, 52, 1055-1066.

Zellner, A., 1976, "Bayesian and Non-Bayesian Analysis of the Regression Model with Multivariate Student-t Error Terms," *Journal of the American Statistical Association*, 71, 400-405.