Limit expressions for the risk of James-Stein estimators*

Abbreviated title: LIMIT EXPRESSIONS

George Casella[1] and Jiunn T. Hwang[2]

Cornell University, Ithaca, New York

Key words and phrases: Multivariate normal, minimax estimation.

AMS 1980 subject classification: Primary 62C99, secondary 62F10.

ABSTRACT

Limit expressions (as $p \to \infty$) are derived for the relative risk of the James-Stein estimator and its positive part version. The limit is simple to evaluate, and gives the amount of improvement in risk that is possible. The technique used is to bound the risk, both above and below, with bounds that converge to the same limit. For the James-Stein estimator these bounds are quite simple to calculate, and are reasonably accurate even for moderate dimensions.

1. INTRODUCTION

It has long been known that the improvement in risk that is obtained by using a James-Stein estimator, instead of the maximum likelihood estimator, increases as the dimension of the problem increases. We investigate the limiting behavior (as $p \to \infty$) of the risk ratio, and obtain a simple expression for the risk improvement.

---

* This paper is BU-766-M in the Biometrics Unit Series.

[1] Research supported by National Science Foundation Grant No. MSC81-02541.

[2] Research supported by National Science Foundation Grant No. MCS80-03568.

The technique used is to obtain upper and lower bounds on the risk ratio, and show that these bounds converge to the same value. These bounds are also useful from a practical point of view, since they can give an experimenter a good idea of the risk improvement. The bounds are quite accurate, even for moderate values of p, and they are easy to compute. Table 1 illustrates the accuracy of the bounds.

A key step in the argument is obtaining bounds on Et^{-1}, where t is a non-central chi-square random variable. Lemma 1, which establishes these bounds, follows quickly from a little known identity, and may be of independent interest.

Let X be an observation from a p-variate normal distribution with mean θ and identity covariance matrix. If $\delta(X)$ is an estimator of θ, define its risk by

$$R[\theta,\delta(X)] = E|\theta - \delta(X)|^2 \quad ,$$

where $|\cdot|$ is the Euclidean norm. The James-Stein estimator

$$\delta^{JS}(X) = \left(1 - \frac{p-2}{|X|^2}\right)X$$

and its positive part version

$$\delta^{+}(X) = \left(1 - \frac{p-2}{|X|^2}\right)^{+} X$$

both dominate X in risk for all θ . The limits (as $p \to \infty$) of the ratios $R[\theta,\delta^{JS}(X)]/R(\theta,X)$ and $R[\theta,\delta^{+}(X)]/R(\theta,X)$ are obtained. In both cases, the domination can be significant.

We start with our preliminary lemma

LEMMA 1: Let $X \sim N(\theta, I)$. If $p \geq 3$, then

$$\frac{1}{p+|\theta|^2} \leq E\,\frac{1}{X'X} \leq \frac{1}{p-2}\left(\frac{p+2}{p+2+|\theta|^2}\right) \quad .$$

Proof: The lower bound follows immediately from Jensen's inequality. To establish the upper bound, use the identity

$$E\{h[\chi_p^2(|\theta|^2)]\chi_p^2(|\theta|^2)\} = pEh[\chi_{p+2}^2(|\theta|^2)] + |\theta|^2 Eh[\chi_{p+4}^2(|\theta|^2)] \quad , \qquad (1.1)$$

where $h(\cdot)$ is a real valued function and $\chi_\nu^2(\lambda)$ denotes a non-central chi-square random variable with ν degrees of freedom and non-centrality parameter $\lambda/2$. Note that $E|X|^2 = E\chi_p^2(|\theta|^2)$. (This identity can be found either in Bock (1975) or Casella (1980).) Now, set $h(y) = y^{-1}$. The identity can now be written as

$$(p-2)E[\chi_p^2(|\theta|^2)]^{-1} + |\theta|^2 E[\chi_{p+2}^2(|\theta|^2)]^{-1} = 1 \quad ,$$

and, after rearranging terms, we obtain

$$E[|X|^2]^{-1} = E[\chi_p^2(|\theta|^2)^{-1} = \frac{1}{p-2}\{1 - |\theta|^2 E[\chi_{p+2}^2(|\theta|^2)]^{-1}\} \quad .$$

Since $E[\chi_{p+2}^2(|\theta|^2)]^{-1} \geq (p+2+|\theta|^2)^{-1}$, by Jensen's inequality, the result follows. Q.E.D.

2. BOUNDS ON THE RISK RATIO

It is now a simple matter, with the help of Lemma 1, to bound the risk ratio $R(\theta,\delta^{JS})/R(\theta,X)$. However, we must also be concerned with the behavior of $|\theta|$ as $p\to\infty$. Since $|\theta|$ depends on p, it would be difficult to interpret any limit expression with some classification of this dependence. The most useful one seems to be to require $\lim_{p\to\infty}(|\theta|^2/p) = c$, for some constant c . The interpretation of the limit of the risk ratio then becomes clear; it is the limiting risk improvement on a θ-sphere of radius c .

THEOREM 1: If $(|\theta|^2/p) \to c$ as $p \to \infty$, then

$$\lim_{p\to\infty} \frac{R(\theta,\delta^{JS})}{R(\theta,X)} = \frac{c}{c+1} \quad .$$

Proof: An integration by parts establishes that

$$\frac{R(\theta,\delta^{JS})}{R(\theta,X)} = 1 - \frac{(p-2)}{p} E|X|^{-2} \quad ,$$

and from Lemma 1 we obtain

$$1 - \frac{p-2}{p}\left(\frac{p+2}{p+2+|\theta|^2}\right) \le \frac{R(\theta,\delta^{JS})}{R(\theta,X)} \le 1 - \frac{(p-2)^2}{p}\left(\frac{1}{p+|\theta|^2}\right) \quad .$$

It is now a simple matter to establish that both the upper and lower bound converge to $c/(c+1)$ as $|\theta|^2/p \to c$ and $p \to \infty$. Q.E.D.

The accuracy of the approximation is quite good. This can be seen by noting that the difference in the bounds is

$$\frac{2(p-2)}{p}\left[\frac{p+2+2|\theta|^2}{(p+|\theta|^2)(p+2+|\theta|^2)}\right] \quad .$$

This quantity, for fixed p, decreases as $|\theta|^{-2}$, thus lower dimensional approximations can be made with reasonable accuracy. For example, if $p = 7$ and $|\theta| = 10$, the difference in the bounds is less than .03 . Table 1 gives the risk ratio $R(\theta,\delta^{JS})/R(\theta,X)$, and the upper and lower bounds, for selected values of p and $|\theta|$. It can be seen that the lower bound is better approximation for smaller $|\theta|$, while the upper bound is better for larger $|\theta|$.

3. THE POSITIVE PART JAMES-STEIN ESTIMATOR

The results for the positive part James-Stein estimator are similar to those for the ordinary James-Stein estimator, although the calculations are a bit more difficult.

THEOREM 2: If $(|\theta|^2/p) \to c$ as $p \to \infty$, then

$$\lim_{p\to\infty} \frac{R(\theta,\delta^+)}{R(\theta,X)} = \frac{c}{c+1} \quad .$$

Proof: Straightforward calculation yields

$$\frac{R(\theta,\delta^+)}{R(\theta,X)} = \frac{|\theta|^2}{p} P_\theta(|X|^2 < p-2) + P(|X|^2 > p-2) - \frac{(p-2)^2}{p} E[g(|X|^2)] \quad , \qquad (3.1)$$

where $g(y) = I_{(y>p-2)}(y)/y$. We now obtain bounds on $Eg(|X|^2)$. Since the indicator function is nondecreasing, and $|X|^{-2}$ is decreasing,

$$\begin{aligned} Eg(|X|^2) &\le P(|X|^2 > p-2) E|X|^{-2} \\ &\le \frac{P(|X|^2 > p-2)}{p-2} \left(\frac{p+2}{p+2+|\theta|^2}\right) \quad , \end{aligned} \qquad (3.2)$$

where the last inequality follows from Lemma 1. Now substituting $g(\cdot)$ for $h(\cdot)$ in (1.1), we have

$$P[\chi^2_p(|\theta|^2) > p-2] = pEg[\chi^2_{p+2}(|\theta|^2)] + |\theta|^2 Eg[\chi^2_{p+4}(|\theta|^2)] \quad ,$$

or, by rearranging terms,

$$Eg(|X|^2) = \frac{1}{p-2}\left\{P[\chi^2_{p-2}(|\theta|^2) > p-2] - |\theta|^2 Eg[\chi^2_{p+2}(|\theta|^2)]\right\}$$

$$\geq \frac{1}{p-2}\left\{P[\chi^2_{p-2}(|\theta|^2) > p-2] - \frac{|\theta|^2}{p} P[\chi^2_{p+2}(|\theta|^2) > p-2]\left(\frac{p+4}{p+4+|\theta|^2}\right)\right\} ,$$

where the last inequality follows from (3.2). Hence we have established upper and lower bounds on $Eg(|X|^2)$. Substituting these bounds into (3.1) and taking limits establishes the theorem. Q.E.D.

REFERENCES

Bock, M. E. (1975). Minimax estimation of the mean of a multivariate normal distribution. Ann. Statist., 3, 209-218.

Casella, G. (1980). Minimax ridge regression estimation. Ann. Statist., 8, 1036-1056.

TABLE 1: Exact values of the risk ratio $R(\theta,\delta^{JS})/R(\theta,X)$ and the upper and lower bounds. The first entry is the lower bound, the middle entry is the exact value, and the third entry is the upper bound.

$\|\theta\|$	p = 3	5	7	9	11	13	15	25
0	.6667 .6667 .8889	.4000 .4000 .6400	.2857 .2857 .4898	.2222 .2222 .3951	.1818 .1818 .3306	.1538 .1538 .2840	.1333 .1333 .2489	.0800 .0800 .1536
1	.7222 .7584 .9167	.4750 .5046 .7000	.3571 .3774 .5536	.2870 .3014 .4556	.2403 .2508 .3864	.2067 .2148 .3352	.1815 .1879 .2958	.1129 .1155 .1862
2	.8148 .8933 .9524	.6182 .6940 .8000	.5055 .5625 .6753	.4296 .4726 .5812	.3743 .4075 .5091	.3320 .3582 .4525	.2984 .3197 .4070	.1987 .2081 .2703
3	.8810 .9563 .9722	.7375 .8262 .8714	.6429 .7181 .7768	.5722 .6338 .6975	.5165 .5671 .6318	.4712 .5131 .5769	.4333 .4686 .5306	.3100 .3273 .3776
4	.9206 .9775 .9825	.8174 .8951 .9143	.7429 .8158 .8447	.6831 .7475 .7822	.6332 .6892 .7273	.5906 .6393 .6790	.5535 .5961 .6366	.4223 .4459 .4839
5	.9444 .9861 .9881	.8687 .9310 .9400	.8109 .8736 .8884	.7623 .8208 .8399	.7201 .7733 .7955	.6827 .7308 .7551	.6492 .6927 .7183	.5223 .5494 .5768
10	.9841 .9966 .9968	.9607 .9822 .9829	.9410 .9654 .9666	.9229 .9482 .9501	.9059 .9313 .9337	.8896 .9148 .9176	.8741 .8988 .9020	.8044 .8258 .8307
15	.9928 .9985 .9985	.9819 .9920 .9820	.9725 .9844 .9846	.9637 .9764 .9767	.9553 .9683 .9688	.9471 .9602 .9609	.9391 .9524 .9531	.9014 .9142 .9154