Bayes Estimates as Expanders
in One and Two Dimensions

by

Anirban DasGupta* and Herman Rubin
Purdue University

Technical Report # 93-38

Department of Statistics
Purdue University

July, 1993

*Research supported by NSF grant DMS 93-077-27.

△

BAYES ESTIMATES AS EXPANDERS
IN ONE AND TWO DIMENSIONS

by

Anirban DasGupta and Herman Rubin
Department of Statistics
Purdue University
1399 Math Sciences Bldg.
West Lafayette, IN 47907-1399

ABSTRACT

Consider the problem of estimating a p-variate normal mean using squared error loss. We demonstrate that contrary to what one might expect,

a $\sup E\left\{\frac{\|d(\underset{\sim}{X})\|}{\|\underset{\sim}{X}\|}\right\} > 1$ for $p \leq 2$, where the supremum is over all priors G symmetric about $\underset{\sim}{0}$, $E(\cdot)$ denotes marginal expectation corresponding to G and $d(\underset{\sim}{X})$ denotes Bayes rule with respect to G.

b The above supremum is equal to 1 for all $p \geq 3$.

Thus Bayes estimates corresponding to symmetric priors can be expanders only in one and two dimensions.

For $p = 1, 2$, the numerical values of the supremum are obtained. We also prove that $\sup E\left\{\frac{\|d(\underset{\sim}{X})\|}{\|\underset{\sim}{X}\|}\right\} = 1$ for every $p > 1$ if one restricts to spherically symmetric priors. In the process, some closed form formulas of independent interest are also obtained.

Key words: Multivariate normal, Bayes estimate, prior, symmetric, marginal, spherically symmetric, supremum.

1. Introduction. Consider the problem of estimating the mean $\underset{\sim}{\theta}$ of a p-variate $N(\underset{\sim}{\theta}, I)$ distribution. If $\underset{\sim}{\theta}$ has a p-variate normal prior with mean $\underset{\sim}{0}$, then the Bayes estimate $d(\underset{\sim}{x})$ of $\underset{\sim}{\theta}$ for squared error loss is a shrinkage estimate in the sense $||d(\underset{\sim}{x})|| \leq ||\underset{\sim}{x}||$ for all $\underset{\sim}{x}$. Similar facts are true for some other common estimation problems, for example for estimating a Binomial success probability using a conjugate Beta prior symmetric about 0.5. It is fair to say that perhaps because of such natural examples, it is quite common to associate Bayes estimates with shrinkage. This article shows that for the problem of estimating a p-variate normal mean using a symmetric prior, Bayes estimates need not completely shrink the unbiased estimate X in low dimensions. It is proved that Bayes estimates can be expanders in the sense that it is possible for the marginal expectation $E\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}$ to be larger than 1 for appropriate symmetric priors in this problem, if $p \leq 2$. For $p = 1$ and 2, we establish the supremum of $E\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}$ over symmetric priors to be finite but larger than 1. We then prove that for each $p \geq 3$, the above supremum is exactly equal to 1. In other words, symmetric prior Bayes estimates can be expanders, but only in less than three dimensions. For each of $p = 1, 2$, we also evaluate the value of the finite supremum, subject to the accuracy of a numerical computation.

Spherically symmetric estimates are quite common in the normal problem; there is a particularly huge literature on properties of spherically symmetric estimates of normal means. See Strawderman (1974), Brandwein and Strawderman (1990), Brandwein (1979), Bock (1985), Cellier et al (1989), among others. Spherically symmetric Bayes estimates, however, can only arise from spherically symmetric priors for the mean vector $\underset{\sim}{\theta}$. It therefore seems quite natural to ask if $E\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}$ can exceed 1 for spherically symmetric priors as well. Since the answer is in the affirmative for symmetric priors in two dimensions as we describe above, it is at least somewhat surprising that we are able to prove that $E\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||} < 1$ for any spherically symmetric prior in dimension $p = 2$. The supremum, however, equals exactly 1. In view of the preceding results for symmetric priors, the case of 3 or more dimensions need not be considered separately for spherically symmetric priors. For the case $p = 1$, we also give a verifiable necessary and sufficient condition on a symmetric prior G for $E\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}$ to exceed 1.

In our view, the combination of the facts: (i) the supremum over symmetric priors is larger than 1 in low dimensions, (ii) there is a clear dimensionality effect and the supremum is 1 precisely from three dimensions, and (iii) with spherically symmetric priors with or without unimodality, the situation reverses to complete shrinkage, is intriguing. These results can be regarded as a new kind of Stein effect. We believe these results are true in similar forms in some other problems. It is not our intention to stimulate much further work on this issue. But we find the combination of the phenomena we establish to be quite surprising. We also hope these results aid in a better understanding of Bayesian inference.

2. A Basic Lemma. The result in this section is simple and yet fundamental for the main results on symmetric priors to follow. It is also completely general in the sense that there is no role of normality in the result.

Lemma 2.1. Let $\underset{\sim}{X} \sim p(\underset{\sim}{x}|\underset{\sim}{\theta})$ and let $\underset{\sim}{\theta}$ have the prior μ. Let $d_\mu(\underset{\sim}{X})$ denote the posterior mean of $\underset{\sim}{\theta}$ corresponding to the prior μ. Then the marginal expectation $E\left\{\frac{||d_\mu(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\}$ is a convex functional of μ.

Proof: Clearly, $E\left\{\frac{||d_\mu(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\}$

$$= \int \frac{|| \int \underset{\sim}{\theta} p(\underset{\sim}{x}|\underset{\sim}{\theta}) d\mu(\underset{\sim}{\theta})||}{||\underset{\sim}{x}||} d\underset{\sim}{x}, \tag{2.1}$$

from which the required convexity follows on using the two facts that integral is a linear operator and $||\cdot||$ is convex, i.e., $||\lambda \underset{\sim}{x} + (1-\lambda)\underset{\sim}{y}|| \leq \lambda||\underset{\sim}{x}|| + (1-\lambda)||\underset{\sim}{y}||$.

Corollary 2.2. Consider the problem of estimating a p-variate normal mean using squared error loss. Let $d(\underset{\sim}{X}) = d_\mu(\underset{\sim}{X})$ denote the Bayes rule for a symmetric prior μ. Then the supremum of $E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\}$ over all symmetric priors equals the supremum of $E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\}$ over all two point symmetric priors.

3. Symmetric Priors. In this section, we will prove the results stated in Section 1 for symmetric priors. For this, we will first derive an explicit formula for $E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\}$ when $d(\underset{\sim}{X})$ is a Bayes rule with respect to a two point symmetric prior supported on $\pm\underset{\sim}{\mu}$. We will establish the supremum of this quantity over $\underset{\sim}{\mu}$ to be 1 for $p \geq 3$ and a finite number larger than 1 for $p = 1, 2$. An approximate numerical value of the supremum will be provided for the cases $p = 1, 2$. The derivation for $p \geq 2$ is formally different from the univariate case. We start with the case $p \geq 2$. It will henceforth be assumed that we are considering estimation of p-variate normal means using squared error loss. It will also be assumed that $d(\underset{\sim}{X})$ denotes Bayes rule and $\mathcal{F}, \mathcal{F}_2$ respectively denote the class of all symmetric and two point symmetric priors. It will also be understood that $E(\cdot)$ stands for marginal expectation.

Theorem 3.1. Let G_μ denote the two point symmetric prior supported at $\pm\underset{\sim}{\mu}$. Then, for $p \geq 2$,

$$E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\} = \frac{\Gamma(\frac{p}{2})}{\sqrt{\pi}\Gamma(\frac{p+1}{2})}||\underset{\sim}{\mu}||^2 e^{-\frac{||\underset{\sim}{\mu}||^2}{2}} {}_2F_2\left(\frac{p}{2}, 1, \frac{p+1}{2}, \frac{3}{2}, \frac{||\underset{\sim}{\mu}||^2}{2}\right), \tag{3.1}$$

where ${}_2F_2$ denotes the generalized hypergeometric series ${}_pF_q$ with $p = q = 2$.

Proof: The proof consists of the following steps.

Step 1. It follows from the definition of $d(\underset{\sim}{x})$ that

$$||d(\underset{\sim}{x})|| = ||\underset{\sim}{\mu}|| \cdot \frac{|e^{-\frac{1}{2}(x-\mu)'(x-\mu)} - e^{-\frac{1}{2}(x+\mu)'(x+\mu)}|}{e^{-\frac{1}{2}(x-\mu)'(x-\mu)} + e^{-\frac{1}{2}(x+\mu)'(x+\mu)}}.$$

Therefore,

$$E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\} = \frac{||\underset{\sim}{\mu}||e^{-\frac{\mu'\mu}{2}}}{2\cdot(2\pi)^{p/2}}\int \frac{e^{-\frac{x'x}{2}}}{||x||}\cdot|e^{\mu'x} - e^{-\mu'x}|dx \tag{3.2}$$

Step 2. Transforming to spherical coordinates $r, u, u_2, \ldots, u_{p-1}$ where $0 \leq r < \infty$ denotes $||\underset{\sim}{x}||$ and $-\frac{\pi}{2} < u \leq \frac{\pi}{2}$ denotes the angle between $\underset{\sim}{\mu}$ and $\underset{\sim}{x}$, (see Anderson (1984), page 279–280) (3.2) gives

$$E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\}$$

$$= K||\underset{\sim}{\mu}||e^{-\frac{\mu'\mu}{2}}\int_0^\infty\int_{-\pi/2}^{\pi/2} e^{-\frac{r^2}{2}}r^{p-2}|e^{r||\mu||sinu} - e^{-r||\mu||sinu}|\overset{p-2}{cos}u\,dudr$$

$\left(\text{where } K = \frac{2^{\frac{p}{2}-2}\Gamma(\frac{p}{2})}{\pi\Gamma(p-1)};\text{ note that } K = \frac{V(p)}{\left(\int_0^1\int_{-\pi/2}^{\pi/2} r^{p-1}cos^{p-2}ududr\right)\cdot 2\cdot(2\pi)^{p/2}},\right.$ where

$V(p)$ is the volume of the unit sphere. See page 246 in Billingsley (1986)$\Big)$

$$= 2K||\underset{\sim}{\mu}||e^{-\frac{\mu'\mu}{2}}\int_0^\infty\int_0^{\pi/2} e^{-\frac{r^2}{2}}r^{p-2}|e^{r||\mu||sinu} - e^{-r||\mu||sinu}|\overset{p-2}{cos}u\,dudr$$

$$= 2K||\underset{\sim}{\mu}||e^{-\frac{\mu'\mu}{2}}\int_0^\infty\int_0^1 |e^{r||\mu||t} - e^{-r||\mu||t}|(1-t^2)^{\frac{p-3}{2}}r^{p-2}e^{-\frac{r^2}{2}}dtdr$$

$$= 2K||\underset{\sim}{\mu}||e^{-\frac{\mu'\mu}{2}}\int_0^\infty\int_0^1 \left(e^{r||\mu||t} - e^{-r||\mu||t}\right)(1-t^2)^{\frac{p-3}{2}}r^{p-2}e^{-\frac{r^2}{2}}dtdr$$

$$= 4K||\underset{\sim}{\mu}||e^{-\frac{\mu'\mu}{2}} \int_0^\infty \sum_{k=0}^\infty \frac{||\underset{\sim}{\mu}||^{2k+1} r^{p+2k-1}}{(2k+1)!} e^{-\frac{r^2}{2}} \left(\int_0^1 t^{2k+1}(1-t^2)^{\frac{p-3}{2}} dt \right) dr, \tag{3.3}$$

where the iterated integration and an interchange of summation and integration are justified by Fubini and the Monotone convergence theorem respectively. Now on using the fact that

$$\int_0^1 t^{2k+1}(1-t^2)^{\frac{p-3}{2}} dt = \frac{1}{2} \frac{\Gamma(k+1)\Gamma(\frac{p-1}{2})}{\Gamma(k+\frac{p+1}{2})}$$

(see Gradshteyn and Ryzhik (1980), page 294, item 3.251) and another interchange of integration and summation, (3.3) reduces to

$$E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\} = \beta||\underset{\sim}{\mu}||e^{-\frac{\mu'\mu}{2}} \sum_{k=0}^\infty \frac{||\underset{\sim}{\mu}||^{2k+1}}{(2k+1)!} \frac{\Gamma(k+1)}{\Gamma(k+\frac{p+1}{2})} \left(\int_0^\infty r^{p+2k-1} e^{-\frac{r^2}{2}} dr \right) \tag{3.4}$$

$$\left(\text{where } \beta = 2K\Gamma\left(\frac{p-1}{2}\right)\right)$$

Making the change of variable $\frac{r^2}{2} = z$ in (3.4), one gets

$$E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\} = \beta||\underset{\sim}{\mu}||e^{-\frac{\mu'\mu}{2}} \sum_{k=0}^\infty \frac{||\underset{\sim}{\mu}||^{2k+1}\Gamma(k+1)2^{k+\frac{p}{2}-1}\Gamma(k+\frac{p}{2})}{(2k+1)!\Gamma(k+\frac{p+1}{2})} \tag{3.5}$$

Step 3. The reduction now is to use Legendre's duplication formula

$$\Gamma(2x) = \frac{2^{2x-1}}{\sqrt{\pi}}\Gamma(x)\Gamma\left(x+\frac{1}{2}\right) \tag{3.6}$$

(see Dettman (1965), page 199) with $x = k+1$. Note that (3.6) gives

$$\frac{\Gamma(k+1)}{(2k+1)!} = \frac{\sqrt{\pi}}{2^{2k+1}\Gamma(k+\frac{3}{2})} \tag{3.7}$$

Use of (3.7) in (3.5) results in

$$E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\} = \beta\sqrt{\pi}||\underset{\sim}{\mu}||e^{-\frac{\mu'\mu}{2}} \sum_{k=0}^\infty \frac{||\underset{\sim}{\mu}||^{2k+1}}{2^{k-\frac{p}{2}+2}} \frac{\Gamma(k+\frac{p}{2})}{\Gamma(k+\frac{3}{2})\Gamma(k+\frac{p+1}{2})} \tag{3.8}$$

(3.8) immediately gives

$$E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\} = \frac{\beta\sqrt{\pi}}{2^{\frac{3-p}{2}}}||\underset{\sim}{\mu}||e^{-\frac{\mu'\mu}{2}} \sum_{k=0}^\infty \left(\frac{||\underset{\sim}{\mu}||}{\sqrt{2}}\right)^{2k+1} \frac{\Gamma(k+\frac{p}{2})}{\Gamma(k+\frac{3}{2})\Gamma(k+\frac{p+1}{2})} \tag{3.9}$$

Step 4. On expanding explicitly the infinite series in (3.9) as $T_0+T_1+T_2+\ldots$ and writing it as $T_0(1+\frac{T_1}{T_0}+\frac{T_2}{T_0}+\ldots)$, and on now using the definition of the constant β, (3.9) results in

$$E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\} = \frac{||\underset{\sim}{\mu}||^2 e^{-\frac{\mu'\mu}{2}}}{2(\Gamma(\frac{3}{2}))^2}\cdot\left[1+\frac{\left(\frac{||\mu||^2}{2}\right)\cdot\frac{p}{2}}{(\frac{3}{2})\cdot\frac{p+1}{2}}+\frac{\left(\frac{||\mu||^2}{2}\right)^2\cdot\frac{p}{2}(\frac{p}{2}+1)}{(\frac{3}{2})(\frac{5}{2})(\frac{p+1}{2})(\frac{p+1}{2}+1)}+\cdots\right]$$

$$= \frac{\Gamma(\frac{p}{2})}{\sqrt{\pi}\Gamma(\frac{p+1}{2})}||\underset{\sim}{\mu}||^2 e^{-\frac{\mu'\mu}{2}}\cdot {}_2F_2\left(\frac{p}{2},1,\frac{p+1}{2},\frac{3}{2},\frac{||\underset{\sim}{\mu}||^2}{2}\right) \tag{3.10}$$

(see Gradshteyn and Ryzhik (1980), page 1045 for the definition of ${}_2F_2$). This finishes the proof of Theorem 3.1.

Corollary 3.2. $\sup\limits_{\mathcal{F}} E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\}$

$$= \frac{2\Gamma(\frac{p}{2})}{\sqrt{\pi}\Gamma(\frac{p+1}{2})}\cdot\sup_{x\geq 0} xe^{-x}\cdot {}_2F_2\left(\frac{p}{2},1,\frac{p+1}{2},\frac{3}{2},x\right) \tag{3.11}$$

Proof: Use Corollary 2.2, Theorem 3.1, and identify $\frac{||\underset{\sim}{\mu}||^2}{2}$ as x.

Next, we will first treat the case $p=2$; this case is technically different from the case $p\geq 3$.

Theorem 3.3. $\sup\limits_{\mathcal{F}} E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\}<\infty$ for $p=2$.

Proof: For this, clearly it is enough to prove that $xe^{-x}{}_2F_2(1,1,\frac{3}{2},\frac{3}{2},x)$ is uniformly bounded for $x\geq 0$. This follows from the definition

$${}_2F_2(x) = 1+\sum_{k=1}^{\infty}\left(\frac{(1)_k}{(\frac{3}{2})_k}\right)^2\frac{x^k}{k!},$$

where

$$
\begin{aligned}
\left(\frac{(1)_k}{(\frac{3}{2})_k}\right)^2 &= \left(\frac{1\times 2\times\ldots\times k}{\frac{3}{2}\times(\frac{3}{2}+1)\times\ldots\times(\frac{3}{2}+k-1)}\right)^2 \\
&= \left(\frac{1}{\frac{3}{2}}\right)^2\cdot\left(\frac{2}{\frac{3}{2}+1}\right)^2\cdots\left(\frac{k}{\frac{3}{2}+k-1}\right)^2 \\
&\leq \frac{1}{2}\cdot\frac{2}{3}\cdot\ldots\frac{k}{k+1}\left(\because\left(\frac{k}{\frac{3}{2}+k-1}\right)^2\leq\frac{k}{k+1}\text{ for all }k\right) \\
&= \frac{1}{k+1} \\
&\therefore\text{ for } x\geq 0, \\
&{}_2F_2\left(1,1,\frac{3}{2},\frac{3}{2},x\right) \\
&\leq 1+\frac{x}{2}+\frac{x^2}{3!}+\frac{x^3}{4!}+\ldots \\
&\Rightarrow x{}_2F_2\left(1,1,\frac{3}{2},\frac{3}{2},x\right)\leq e^x-1 \\
&\Rightarrow xe^{-x}{}_2F_2\left(1,1,\frac{3}{2},\frac{3}{2},x\right)\leq 1-e^{-x}\leq 1
\end{aligned}
$$

proving the required result.

Discussion. In spite of Theorem 3.3, the only possible way to evaluate the value of the finite supremum of $E\left\{\frac{\|d(\underset{\sim}{X})\|}{\|\underset{\sim}{X}\|}\right\}$ for $p=2$ seems to be a numerical maximization of (3.11) for $p=2$. This is because in contrast to the case $p\geq 3$ we will treat shortly, for $p=2$ the maxma of (3.11) is an internal maxima. In our experience, the computational reliability when dealing with an infinite series is less than the reliability for computing integrals on bounded sets. We therefore give below an equivalent integral representation of (3.11). The proof follows on using the following representations.

Lemma 3.4.

$$
{}_2F_2(\beta,\alpha_2,\beta+\sigma,\rho_2,x)=\frac{\Gamma(\beta+\sigma)}{\Gamma(\beta)\Gamma(\sigma)}\int_0^1 t^{\beta-1}(1-t)^{\sigma-1}{}_1F_1(\alpha_2,\rho_2,xt)dt, \tag{3.12}
$$

and

$$
{}_1F_1(a,c,z)=\frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)}\int_0^1 e^{zs}s^{a-1}(1-s)^{c-a-1}ds, \tag{3.13}
$$

if $c>a>0$.

For (3.12), see Luke (1975), page 161, and for (3.13), see page 284 of the same reference.

Theorem 3.5. For $p = 2$,

$$\infty > \sup_{G \in \mathcal{F}} E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\} = \frac{1}{\pi} \cdot \sup_{x \geq 0} x \cdot \int_0^1 \int_0^1 \frac{e^{xst - x(s+t)}}{\sqrt{st}} ds dt$$
$$= 1.06691 \text{ (approximately).}$$

We will next demonstrate that for $p \geq 3$, the supremum we seek is equal to 1. We will achieve most of what is required to prove this in the form of a Lemma.

Lemma 3.6. For every $p \geq 3$, the function $xe^{-x}{}_2F_2\left(\frac{p}{2}, 1, \frac{p+1}{2}, \frac{3}{2}, x\right)$ is monotone nondecreasing in x, for $x \geq 0$.

Proof: Write

$$\begin{aligned} & xe^{-x}{}_2F_2\left(\frac{p}{2}, 1, \frac{p+1}{2}, \frac{3}{2}, x\right) \\ = \; & \frac{{}_2F_2\left(\frac{p}{2}, 1, \frac{p+1}{2}, \frac{3}{2}, x\right)}{\frac{e^x - 1}{x}} \cdot (1 - e^{-x}) \end{aligned} \tag{3.14}$$

Since $1 - e^{-x}$ is nonnegative and increasing, it will suffice to show that the ratio $\frac{{}_2F_2\left(\frac{p}{2}, 1, \frac{p+1}{2}, \frac{3}{2}, x\right)}{\frac{e^x-1}{x}} \geq 0$ is monotone nondecreasing.

For this, write

$${}_2F_2\left(\frac{p}{2}, 1, \frac{p+1}{2}, \frac{3}{2}, x\right) = \sum_{k=0}^{\infty} a_k x^k$$
$$\text{and } \frac{e^x - 1}{x} = \sum_{k=0}^{\infty} b_k x^k;$$

here $a_k = \frac{(\frac{p}{2})_k (1)_k}{(\frac{p+1}{2})_k (\frac{3}{2})_k} \cdot \frac{1}{k!}$ and $b_k = \frac{1}{(k+1)!}$. It is then very easy to verify that if $p \geq 3$, then the coefficients $\{a_k\}, \{b_k\}$ satisfy

$$\frac{\frac{a_{k+1}}{a_k}}{\frac{b_{k+1}}{b_k}} = \frac{(\frac{p}{2} + k)(k+2)}{(\frac{p+1}{2} + k)(k + \frac{3}{2})} \geq 1 \text{ for all } k,$$

i.e., $\frac{a_k}{b_k}$ is nondecreasing in k. A standard monotone likelihood ratio argument then implies the ratio under consideration is nondecreasing.

Therefore, we have now proved the following theorem.

Theorem 3.7. For every $p \geq 3$,

$$\sup_{G \in \mathcal{F}} E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\} = \frac{2\Gamma(\frac{p}{2})}{\sqrt{\pi}\Gamma(\frac{p+1}{2})} \cdot \lim_{x \to \infty} xe^{-x}{}_2F_2\left(\frac{p}{2}, 1, \frac{p+1}{2}, \frac{3}{2}, x\right). \tag{3.15}$$

We will now demonstrate that (3.15) equals 1 for each $p \geq 3$.

Theorem 3.8. For each $p \geq 3$, $\sup_{G \in \mathcal{F}} E\left\{\frac{\|d(\underset{\sim}{X})\|}{\|\underset{\sim}{X}\|}\right\} = 1$.

Proof: By virtue of Theorem 3.7, it is enough to show that

$$\lim_{x \to \infty} x e^{-x} {}_2F_2\left(\frac{p}{2}, 1, \frac{p+1}{2}, \frac{3}{2}, x\right) = \frac{\sqrt{\pi}\Gamma(\frac{p+1}{2})}{2\Gamma(\frac{p}{2})}.$$

Since $1 - e^{-x} \to 1$ as $x \to \infty$, it will suffice to show that

$$\lim_{x \to \infty} \frac{{}_2F_2\left(\frac{p}{2}, 1, \frac{p+1}{2}, \frac{3}{2}, x\right)}{\frac{e^x - 1}{x}} = \frac{\sqrt{\pi}\Gamma(\frac{p+1}{2})}{2\Gamma(\frac{p}{2})}. \tag{3.16}$$

However, using the notations of Lemma 3.6,

$$\begin{aligned} & \lim_{x \to \infty} \frac{{}_2F_2(\frac{p}{2}, 1, \frac{p+1}{2}, \frac{3}{2}, x)}{\frac{e^x - 1}{x}} \\ = & \lim_{k \to \infty} \frac{a_k}{b_k} \\ = & \lim_{k \to \infty} \frac{(\frac{p}{2})_k (k+1)!}{(\frac{p+1}{2})_k \cdot (\frac{3}{2})_k} \\ = & \lim_{k \to \infty} \frac{\Gamma(\frac{p+1}{2})\Gamma(\frac{3}{2})}{\Gamma(\frac{p}{2})} \cdot \frac{\Gamma(\frac{p}{2} + k) \cdot (k+1)!}{\Gamma(\frac{p+1}{2} + k)\Gamma(\frac{3}{2} + k)} \\ = & \frac{\sqrt{\pi}\Gamma(\frac{p+1}{2})}{2\Gamma(\frac{p}{2})} \cdot \lim_{k \to \infty} \frac{\Gamma(\frac{p}{2} + k)(k+1)!}{\Gamma(\frac{p+1}{2} + k)\Gamma(\frac{3}{2} + k)} \end{aligned} \tag{3.17}$$

Since

$$\lim_{k \to \infty} \frac{\Gamma(\frac{p}{2} + k)(k+1)!}{\Gamma(\frac{p+1}{2} + k)\Gamma(\frac{3}{2} + k)} = 1$$

by Stirling's formula, the theorem follows from (3.17).

It remains to settle the case $p = 1$ now. First we give an analog of Theorem 3.1 for $p = 1$.

Theorem 3.9. Let G_μ denote the two point symmetric prior supported at $\pm\mu$. Then,

$$E\left\{\frac{|d(X)|}{|X|}\right\} = \mu e^{-\frac{\mu^2}{2}} \left(\int_0^\mu e^{\frac{t^2}{2}} dt\right). \tag{3.18}$$

Proof: The derivation mostly follows the lines of Theorem 3.1, but we do not make transformations to spherical coordinates now (because we are in one dimension). We will omit the derivation; it is straightforward. The facts that (3.18) remains bounded but has a maximum larger than 1 need to be demonstrated, though. The next result does this.

Theorem 3.10. $1 < \sup_{G \in \mathcal{F}} E\left\{\frac{|d(X)|}{|X|}\right\} < \infty$ for $p = 1$; its numerical value equals 1.28475 (approximately).

Proof: To prove that the supremum is finite, simply consider the function

$$g(\mu) = \frac{\mu \int_0^\mu e^{t^2/2} dt}{e^{\mu^2/2}} \tag{3.19}$$

Two applications of L'Hospital's rule show that $g(\mu) \to 1$ as $\mu \to \infty$; since $g(0) = 0$ and g is continuous, it follows g is bounded and therefore so is $E\left\{\frac{|d(X)|}{|X|}\right\}$ by virtue of Theorem 3.9.

To prove that the supremum is larger than 1, we will prove $g(2) > 1$. This is easily done since

$$\begin{aligned} g(2) &= \frac{2}{e^2} \int_0^2 e^{t^2/2} dt \\ &> \frac{2}{e^2} \cdot \int_0^2 \left(1 + \frac{t^2}{2} + \frac{t^4}{8}\right) dt > 1.11 > 1. \end{aligned}$$

The numerical value 1.28475 is obtained on numerical maximization of (3.18) over $\mu \geq 0$.

Even though we have used only two point priors for the technical purpose of obtaining the suprema in all dimensions, we like to point out that absolutely continuous priors can be used as well to arrive at each phenomenon described in this article. For the case $p = 1$, the following result provides a necessary and sufficient condition on a symmetric prior G for the corresponding value of $E\left\{\frac{|d(X)|}{|X|}\right\}$ to exceed 1.

Theorem 3.11. For $p = 1$, let G be any symmetric prior for θ. Then $E\frac{|d(X)|}{|X|} > 1$ if and only if

$$\int_0^\infty \theta e^{-\frac{\theta^2}{2}} \left(\int_0^\theta e^{\frac{x^2}{2}} dx\right) dG(\theta) > \frac{1}{2}. \tag{3.20}$$

Remark. This is a condition that is apparently easily verifiable for any given G. Two point distributions supported on $\pm\theta$ with $\theta \geq 1.32$ satisfy (3.20) and so does any symmetric G supported on $\mathbf{R} - (-1.32, 1.32)$. Since normal distributions can be used to approximate point masses (in the sense of weak convergence), it follows from Helly's theorem that mixtures of (two) normal distributions can also be used to obtain (3.20). The broad message is that many types of symmetric priors result in $E\left\{\frac{|d(X)|}{|X|}\right\} > 1$ for $p = 1, 2$.

Proof of Theorem 3.11: The proof of this theorem is exactly the same as that of Theorem 3.9 and will be omitted. In fact, knowing the statement of Theorem 3.11 makes the assertion of Theorem 3.9 transparent.

4. Spherically Symmetric Priors. In view of the results we obtained in Section 3, it is not unreasonable to expect that Bayes estimates may be expanders for spherically symmetric priors as well in two dimensions. For the sake of completeness, we will prove in this section that this is not the case, i.e., for any spherically symmetric prior in dimension $p = 2, E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\} < 1$. The case $p = 1$, of course, is already treated in Section 3. Once it is known that the supremum cannot exceed 1 with spherically symmetric priors, it is easy to see that the failure to approach two point distributions via spherically symmetric ones is the reason behind this phenomenon.

For the rest of this section, u will denote $||\underset{\sim}{x}||$ and $t = u^2$ will denote $||\underset{\sim}{x}||^2$. Also $I_\nu(z)$ will denote the Bessel I_ν function defined as

$$I_\nu(z) = \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(k+\nu+1)} \left(\frac{z}{2}\right)^{\nu+2k} \tag{4.1}$$

Theorem 4.1. Let $\underset{\sim}{\theta}$ have a spherically symmetric prior distribution G. Then $E\left\{\frac{||d(\underset{\sim}{X})||}{||\underset{\sim}{X}||}\right\}$ ≤ 1 for $p = 2$.

We will prove Theorem 4.1 by breaking it up in a series of Lemmas. For ease of writing the proof, we will present the proof for the case when $\underset{\sim}{\theta}$ has a density $g(||\underset{\sim}{\theta}||)$.

Lemma 4.2. The marginal density $f(\underset{\sim}{x})$ of $\underset{\sim}{x}$ is a function of t alone and is given by

$$f(\underset{\sim}{x}) = m(t) = ke^{-\frac{t}{2}} \int_0^\infty I_0(r\sqrt{t})re^{-\frac{r^2}{2}}dG(r), \tag{4.2}$$

where $k > 0$ is a constant.

Proof: By definition,

$$\begin{aligned} f(\underset{\sim}{x}) &\propto \int e^{-\frac{1}{2}(x-\theta)'(x-\theta)} g(||\underset{\sim}{\theta}||)d\underset{\sim}{\theta} \\ &\propto e^{-\frac{x'x}{2}} \int e^{\theta' x} e^{-\frac{\theta'\theta}{2}} g(||\underset{\sim}{\theta}||)d\underset{\sim}{\theta} \\ &\propto e^{-\frac{x'x}{2}} \int_0^\infty \left(\int_{-\pi}^{\pi} e^{r||\underset{\sim}{x}||\cos u} du\right) e^{-\frac{r^2}{2}} rg(r)dr \end{aligned}$$

(by using spherical coordinates)

$$\propto e^{-\frac{x'x}{2}} \int_0^\infty \left(\int_{-1}^{1} \frac{e^{r\|\underset{\sim}{x}\| z}}{\sqrt{1-z^2}} dz \right) e^{-\frac{r^2}{2}} r g(r) dr$$

$$\propto e^{-\frac{t}{2}} \int_0^\infty I_0(r\sqrt{t}) e^{-\frac{r^2}{2}} r g(r) dr, \tag{4.3}$$

where (4.3) follows on using the integral formula

$$I_0(x) = \frac{1}{\pi} \int_{-1}^{1} \frac{e^{xz}}{\sqrt{1-z^2}} dz$$

(Gradshteyn and Ryzhik (1980), page 958).

Lemma 4.3. The Bayes estimate $d(\underset{\sim}{x})$ is given by

$$d(\underset{\sim}{x}) = \underset{\sim}{x} \left(1 + 2\frac{m'(t)}{m(t)} \right).$$

Proof: This follows immediately on using the well known formula

$$d(\underset{\sim}{x}) = \underset{\sim}{x} + \frac{\nabla f(\underset{\sim}{x})}{f(\underset{\sim}{x})}$$

(see Brown and Hwang (1982)) and then using $f(\underset{\sim}{x}) = m(t)$.

Lemma 4.4. $\frac{\|d(\underset{\sim}{x})\|}{\|\underset{\sim}{x}\|} = 1 + 2\frac{m'(t)}{m(t)}$.

Proof: In view of Lemma 4.3, we have to only prove that $|1 + 2\frac{m'(t)}{m(t)}| = 1 + 2\frac{m'(t)}{m(t)}$, i.e., $1 + 2\frac{m'(t)}{m(t)} \geq 0$ (this just says the Bayes estimate is a nonnegative multiple of $\underset{\sim}{x}$). To prove this, first observe that

$$m'(t) \propto -\frac{1}{2} e^{-\frac{t}{2}} \int_0^\infty I_0(r\sqrt{t}) r e^{-\frac{1}{2}r^2} g(r) dr$$

$$+ \frac{e^{-\frac{t}{2}}}{2\sqrt{t}} \int_0^\infty I_1(r\sqrt{t}) r^2 e^{-\frac{1}{2}r^2} g(r) dr, \tag{4.4}$$

with the same constant of proportionality k as in (4.2); note that in the above we have used the fact $\frac{d}{dz} I_0(z) = I_1(z)$. From (4.4), we immediately have

$$1 + 2\frac{m'(t)}{m(t)} = \frac{\int_0^\infty I_1(r\sqrt{t}) r^2 e^{-\frac{r^2}{2}} g(r) dr}{\sqrt{t} \int_0^\infty I_0(r\sqrt{t}) r e^{-\frac{r^2}{2}} g(r) dr} \geq 0, \tag{4.5}$$

since $I_\nu(z) \geq 0$ for $z \geq 0, \nu = 0, 1$.

Lemma 4.5. $E\left\{\frac{\|d(\underset{\sim}{X})\|}{\|\underset{\sim}{X}\|}\right\} > 1$ iff $\int\limits_0^\infty m'(u^2)u\,du > 0$.

Proof: Clearly, by virtue of Lemma 4.4, $E\left\{\frac{\|d(\underset{\sim}{X})\|}{\|\underset{\sim}{X}\|}\right\} > 1$ iff $E\left\{\frac{m'(t)}{m(t)}\right\} > 0$, from which the result follows on transformation to spherical coordinates.

Lemma 4.6. $\int\limits_0^\infty m'(u^2)u\,du$

$$\propto \int\limits_0^\infty re^{-\frac{r^2}{2}}g(r)\int\limits_0^\infty \left\{rI_1(ru)e^{-\frac{u^2}{2}}\right\}du\,dr$$
$$-\int\limits_0^\infty re^{-\frac{r^2}{2}}g(r)\int\limits_0^\infty \left\{I_0(ru)e^{-\frac{u^2}{2}}u\right\}du\,dr$$

Proof: This is immediate from (4.4) and a separate application of Fubini's theorem to each of the two terms in (4.4).

Lemma 4.7. For each $r \geq 0$,

$$\int\limits_0^\infty rI_1(ru)e^{-\frac{u^2}{2}}du \leq \int\limits_0^\infty I_0(ru)e^{-\frac{u^2}{2}}u\,du$$

Proof: Since $\frac{d}{du}(I_0(ru)) = rI_1(ru)$, integration by parts gives

$$\begin{aligned}&\int\limits_0^\infty rI_1(ru)e^{-\frac{u^2}{2}}du \\ &= I_0(ru)e^{-\frac{u^2}{2}}\Big|_0^\infty + \int\limits_0^\infty I_0(ru)ue^{-\frac{u^2}{2}}du\end{aligned} \tag{4.6}$$

The Lemma follows on noting $I_0(ru)e^{-\frac{u^2}{2}}\Big|_0^\infty = -1$.

Proof of Theorem 4.1: Combine Lemma 4.5, 4.6 and 4.7.

Corollary 4.8. Let $\mathcal{F}_s$ denote the family of spherically symmetric priors for $p = 2$. Then $\sup\limits_{G\in\mathcal{F}_s} E\left\{\frac{\|d(\underset{\sim}{X})\|}{\|\underset{\sim}{X}\|}\right\} = 1$.

Proof: Use Theorem 4.1 and the fact that $E\left\{\frac{\|d(\underset{\sim}{X})\|}{\|\underset{\sim}{X}\|}\right\} = \frac{\tau^2}{1+\tau^2}$ for $G = N(\underset{\sim}{0}, \tau^2 I)$ and $\frac{\tau^2}{1+\tau^2} \to 1$ as $\tau \to \infty$.

5. Summary. The results we describe can be regarded as a new manifestation of Stein effect. The surprise in the results obtained hopefully contribute to better scientific understanding of Bayesian inference.

Acknowledgement. Our original proof for Theorem 3.3 was replaced by the current proof given by Brani Vidakovic. Bill Studden provided formula (3.18). Larry Brown and Bill Strawderman made some insightful comments. We are indebted to all of them.

References

Anderson, T.W. (1984). *An Introduction to Multivariate Statistical Analysis*, John Wiley, New York.

Billingsley, P. (1986). *Probability and Measure*, 2nd Edition, John Wiley, New York.

Bock, M.E. (1985). Minimax estimators that shift towards a hypersphere for location vectors of spherically symmetric distributions, *Jour. Mult. Analysis*, **17**, 127–147.

Brandwein, A.C. (1979). Minimax estimation of the mean of spherically symmetric distributions under quadratic loss, *Jour. Mult. Analysis* , **9**, 579–588.

Brandwein, A.C. and Strawderman, W.E. (1990). Stein estimation: The spherically symmetric case, *Stat. Sci.*, **5**, 356–369.

Brown, L.D. and Hwang, G.J.T. (1982). A unified admissibility proof, *Stat. Decision Theory and Rel. Topics, Proc. of the Fourth Purdue Symposium*, J. Berger and S.S. Gupta, eds., Academic Press, New York.

Cellier, D., Fourdrinier, D. and Robert, C. (1989). Robust shrinkage estimators of location parameters for elliptically symmetric distributions, *Jour. Mult. Analysis*, **29**, 39–52.

Dettman, J.W. (1965). *Applied Complex Variables*, Dover, New York.

Gradshteyn, I.S. and Ryzhik, I.M. (1980). *Tables of Integrals, Series, and Products*, Academic Press, New York.

Luke, Y.L. (1975). *Mathematical Functions and Their Approximations*, Academic Press, New York.

Strawderman, W.E. (1974). Minimax estimation of location parameters for certain spherically symmetric distributions, *Jour. Mult. Analysis*, **4**, 255–264.