Statistical Science
2019, Vol. 34, No. 4, 657–668
https://doi.org/10.1214/19-STS744

Larry Brown's Work on Admissibility

Iain M. Johnstone

Abstract. Many papers in the early part of Brown's career focused on the admissibility or otherwise of estimators of a vector parameter. He established that inadmissibility of invariant estimators in three and higher dimensions is a general phenomenon, and found deep and beautiful connections between admissibility and other areas of mathematics. This review touches on several of his major contributions, with a focus on his celebrated 1971 paper connecting admissibility, recurrence and elliptic partial differential equations.

Key words and phrases: Complete class theorems, inadmissibility, Blyth's method, elliptic partial differential equation, recurrence, Brownian diffusion, loss function, best invariant estimator, James–Stein estimator, variational problem, differential inequality.

Larry Brown was the grandmaster of admissibility. If one includes the work on complete classes, he wrote well over 30 papers, at least, starting with his Ph.D. thesis in 1964. The majority, and the best known, were in the early part of his career, but he continued to write from time to time on topics related to admissibility until recently. He developed the most general rigorous theory, he had powerful heuristics, he used, and elaborated and extended all the methods. He knew concrete cases in detail, including weird and wonderful counterexamples.

Two decades into the next century, admissibility is not on everyone's tongue, even within statistical theory, but Brown's work is part of our enduring heritage, with important and ongoing influence on methodology. For a first example, he showed that there is a general distinction between estimation in lower and higher dimensions, and that it is what we now call a universal phenomenon—not an artifact of which distribution or which loss function is used. As a second example, in concrete special cases, he found beautiful and precise mathematical connections, as in his most famous single admissibility work, the 1971 paper. The sharp results of the 1971 paper are a touchstone for modern work on shrinkage via Bayesian hierarchical models.

Iain M. Johnstone is Professor of Statistics and of Biomedical Data Science, Department of Statistics, Stanford University, Sequoia Hall, 390 Jane Stanford Way, Stanford, California 94305-4020, USA (e-mail: imj@stanford.edu).

This article will attempt a sampling of Brown's comprehensive contributions to admissibility, aiming to illustrate the two points just made (especially the 1971 paper) and to give a taste of some of the other major papers. In addition, an attempt has been made to at least mention most of the published papers on admissibility and to include them in the bibliography. Berger (1985), Chapter 8, contains an extensive review of complete class theorems and admissibility with substantial coverage of Brown's work up to that time. The panoramic survey by Berger and DasGupta (2019) in this volume contains additional discussion and perspective on many of the papers discussed here, and much more.

1. ADMISSIBILITY AND BEST INVARIANT ESTIMATORS

Brown's work on admissibility began with his Ph.D. thesis, written at Cornell under what he describes—in the informative conversation DasGupta (2005)—as the "perceptive but nondirective" guidance of Jack Kiefer. He goes on to say,

> He [Kiefer] told me that Stein was doing some really interesting work on admissibility and I should take a look at that. Statistics was lovely in those days; I essentially had to read five papers to know all the necessary background.

Those five papers are Blackwell (1951), Hodges and Lehmann (1951), Stein (1956, 1959) and James and

Stein (1961), and they set a remarkable standard for interest, originality and concision.

Background. A significant focus of these papers is on the location parameter estimation problem. The setup assumes n independent observations $X_1, \dots, X_n \in \mathbb{R}^p$ from a location family with density $f(x-\theta)$. It is desired to estimate θ using a loss function $L(\theta, d) = W(d-\theta)$. The problem is conventionally transformed to $X = X_1$ and the ancillary statistics $Y_i = X_i - X_1$, with $Y = (Y_2, \dots, Y_n)$. After a further (harmless) transformation $X \to X + g(Y)$ for suitable g, it can be assumed that the best invariant estimator is given by $\delta_0(x, y) = x$.

Stein had shown that for squared error loss, the best invariant estimator was admissible for $p = 1$ (Stein, 1959) and for $p = 2$ (James and Stein, 1961), in both cases by variants of Blyth's method. After fruitless attempts to extend the result to $p \geq 3$, he shocked the statistical world by proving that, on the contrary, at least for X_i from the multivariate normal $N_p(\theta, I)$ and with squared error loss, that $\bar{X}$ was inadmissible. He showed that a dominating estimator could be found of the form $\delta(X) = (1 - b/(a + |X|^2))X$ for b small and a large. In James and Stein (1961), explicit calculations were done for the celebrated James–Stein estimator $\delta(X) = (1 - (p-2)/|X|^2)X$ which showed that very substantial gains in mean squared error could be obtained.

Against this background, Brown set out to develop an encompassing theory. As he told DasGupta,

> My overall goal was to show that what Stein did was not a particular feature of squared error loss or normal distributions, and that indeed there was a very general dichotomy, with something happening in one and two dimensions and the contrary in three or more dimensions.

The results appear in the major paper Brown (1966). In dimension one, Brown extends the approach of Blackwell (1951) for discrete problems to show that the best invariant estimator is admissible in great generality. The result is proven under mild moment-like conditions on $p(x, y)$, the conditional density of X given Y, under growth and regularity conditions on the loss function W, and under a condition guaranteeing uniqueness of the best invariant estimator [to rule out counterexamples such as the one given by Blackwell (1951)]. He even considers sequential problems, in which the loss function can include the cost of observation, for example, $W(t, n) = W_1(t) + W_2(n)$.

As one illustration of the results, in the Gaussian case with standard normal density $\phi(x-\theta)$, the sample mean $\bar{X}$ is admissible if $W(t)$ is nondecreasing as t moves away from 0 in either direction and satisfies the growth condition

$$\int |t| W(2|t|)\phi(t)\, dt < \infty,$$

with the only exceptions being the loss functions $W(0) = a < b \equiv W(t)$ for $t \neq 0$.

As another example, for squared error loss $W(t) = t^2$ and general location densities, Brown's method yields admissibility of $\delta_0(x) = x$ under the condition $\mathbb{E}X^3 < \infty$, which is only slightly weaker than the result of Stein (1959) obtained by a different technique (building on the method of Blyth (1951)), and specifically focused on the quadratic loss case.

On the inadmissibility side, for $p \geq 3$, the main result of Brown (1966) is Theorem 3.1.1, which considers the fixed sample size problem with convex loss functions W. In addition to moment conditions, the main assumption is that the $p \times p$ matrix

$$Q = \mathbb{E}_0[X \nabla W(X)^T / 2]$$

is nonsingular. This will hold true for pth power losses such as $W(t) = \sum |t_i|^p$ or $(\sum |t_i|^p)^{1/p}$ for $p \geq 1$, including $\max_i |t_i|$, but will fail, say, for $W(t) = t_1^2$ or $(\sum t_i)^2$. Brown shows that for $p \geq 3$,

$$\delta(X) = \left(I - \frac{b}{a + |X|^2} Q^{-1}\right) X \tag{1}$$

dominates $\delta_0(X) = X$ for some positive functions $a = a(Y)$ sufficiently large and $b = b(Y)$ sufficiently small.

The estimator is written here in the form (1) to emphasize that it generalizes that used by Stein (1956), where X is $N(\theta, I)$ and $W(t) = |t|^2$ so that $Q = I_p$. Indeed, Brown shows that the Stein's Taylor series argument can be pushed to work in this much more general setting, and that the resulting risk difference bound has the form

$$R(\theta, \delta_0) - R(\theta, \delta) \geq \frac{2b}{a + |\theta|^2}(p - 2 - 2pkb) + o\left(\frac{1}{a + |\theta|^2}\right),$$

with constant k an explicit function of p, Q, and W. This is exactly the form obtained by Stein (with $1/2$ in place of $2pk$ in his (22)), and it follows that δ dominates δ_0 for small b and large a.

The proof that the best invariant estimator is generically admissible in dimension $p = 2$ was deferred to a paper with Martin Fox (Brown and Fox, 1974a). An indication that $p = 2$ is harder appears already in the Gaussian case with squared error, in which the first proof of admissibility of $\delta_0(x) = x$ appears in Section 4 of Stein (1956).

A second paper with Fox, Brown and Fox (1974b), shows how the formulation of Brown (1966) can be extended to show admissibility of the best invariant *decision* procedure in a variety of problems, including testing and confidence sets, that involve an unknown location or scale parameter.

Much later, with Linda Zhao, Brown returned to the heuristic geometrical argument given by Stein (1956) to suggest why the usual estimator should be inadmissible if the dimension is sufficiently large. Brown and Zhao (2012) develops the argument further—by exploiting the spherical symmetry they conceptualize the multidimensional setting into a two-dimensional framework that can be geometrically analyzed.

When $L(\theta, d)$ is not of the form $\rho(d - \theta)$, the clear distinction between $p \le 2$ and $p \ge 3$ can vanish. Stimulated by an example of Berger (1980), Brown (1980a) considered, in the usual $X \sim N_p(\theta, I)$ setting, *weighted coordinate loss functions* of the form

$$L(\theta, a) = \left[\sum_{i=1}^{p} v(\theta_i)\right]^{-1} \sum_{i=1}^{p} v(\theta_i)(\theta_i - a_i)^2.$$

Examples considered include $v(t) = e^{rt}$ and $v(t) = (1 + t^2)^{r/2}$. For the second example, Brown shows that $\delta_0(x) = x$ is inadmissible for $p > (2 - r)/(1 - r)$ and admissible for $p < (2 - r)/(1 - r)$. In particular, δ_0 is admissible *for all* $p \ge 1$, when $r \ge 1$.

1968 scale parameter paper. Brown (1968), another influential work, continues the study of best invariant estimators, this time focused on estimating (any power of) a scale parameter when there is an unknown location parameter. Again the inspiration comes from a concrete result of Stein (1964): for normally distributed observations and squared error, a "pretest" scale-invariant estimator dominates the best location-scale invariant rule. Brown's object is to show that a similar phenomenon holds for a wider class of distributions and quite general invariant loss functions: rather than the best invariant estimator, one should use the usual estimator of σ^2 when $|\bar{x}|/s$ is large, and a somewhat smaller multiple when $|\bar{x}|/s$ is small.

The paper gives examples in which the best invariant estimate of σ^2 is inadmissible in the presence of nuisance parameters, when the corresponding estimate of σ^2 with known values of the nuisance parameters is admissible. It also provides a new justification for Stein's loss function $L(y) = y - 1 - \log y$: it is essentially the only loss function for which the best invariant estimator of a scale parameter is always unbiased.

2. BROWN 1971

The celebrated paper Brown (1971) describes a mathematical characterization of admissibility in terms of recurrence and insoluble boundary value problems. It begins by recalling three striking phenomena:

- in statistics, the best invariant estimator $\delta(x) = x$ from $N_p(\theta, I)$ is admissible if $p \le 2$ and inadmissible if $p \ge 3$ (Stein, 1956),
- in probability, Brownian motion B_t on $\mathbb{R}^p$ is recurrent if $p \le 2$ and transient if $p \ge 3$ (Lévy, 1940, Kakutani, 1944), and
- in differential equations, the classical fact that the exterior Dirichlet problem

$$\Delta u = 0, \quad |x| > 1, \quad u = \begin{cases} 1 & |x| = 1, \\ 0 & |x| \to \infty, \end{cases}$$

has no solution if and only if $p \le 2$. [Here, $\Delta = \nabla \cdot \nabla = \sum_1^p \partial^2 / \partial x_i^2$ is the Laplacian operator.]

Brown shows that there is a close mathematical connection between these phenomena that extends well beyond these invariant cases. He characterizes the admissibility of a generalized Bayes estimator $\delta_F(x)$ based on a prior measure F in terms of the recurrence of an associated diffusion and the nonsolvability of a variational problem and its concomitant exterior Dirichlet problem.

The story begins with the characterization of admissibility in terms of Bayes rules. Every unique proper Bayes rule is admissible, and starting with Wald, it was shown that quite generally, every admissible rule is a limit of proper Bayes rules in an appropriate sense. [Rukhin (1995) gives a helpful survey of admissibility.] The special features of the multivariate normal mean setting under squared error loss allows these results to take concrete form and Brown's paper masterfully exploits the resulting rich structure.

If $G(d\theta)$ is a nonnegative (prior) measure, not necessarily proper, then the marginal density of $X \sim N_p(\theta, I)$ is given by a convolution with the Gaussian density $\phi(x) = (2\pi)^{-p/2} \exp(-|x|^2/2)$,

$$g^*(x) = \phi \star G(x) = \int \phi(x - \theta) G(d\theta).$$

The convolution smooths G so that $g^*(x)$ is typically an analytic function and the posterior mean

$$\delta_G(x) = \frac{\int \theta\phi(x-\theta)G(d\theta)}{\int \phi(x-\theta)G(d\theta)} \tag{2}$$

is called a generalized Bayes estimator (for squared error loss). It can be rewritten, using $\nabla\phi(x) = -x\phi(x)$, in a form that is central to Brown's analysis[1]

$$\delta_G(x) = x + \frac{\nabla g^*(x)}{g^*(x)}. \tag{3}$$

Bayes geometry. When squared error loss $L(\theta, d) = |\theta - d|^2$ is used, with risk function $R(\theta, \delta) = \mathbb{E}_\theta L(\theta, \delta(X))$, the posterior mean (2), (3) minimizes the integrated risk

$$B(G, \delta) = \int R(\theta, \delta)G(d\theta).$$

The *Bayes regret* of δ then has a concrete representation for squared error loss (James and Stein, 1961; Brown, 1971, (1.3.2)):

$$\begin{aligned} &B(G,\delta) - B(G,\delta_G) \\ &\quad = \int \left|\delta(x) - \delta_G(x)\right|^2 g^*(x)\,dx. \end{aligned} \tag{4}$$

Admissibility and Bayes approximation. The Bayes regret appears in the abstract characterization of admissibility of Stein (1955) and was further developed by Farrell (1968a, 1968b). To state it, let $B = \{x \in \mathbb{R}^p : |x| \le 1\}$ be the unit ball. As formulated by Brown, the characterization says that δ is admissible if and only if there exist finite measures G_n with (i) compact support, and (ii) $G_n(B) \ge 1$, so that

$$B(G_n, \delta) - B(G_n, \delta_{G_n}) \to 0. \tag{5}$$

The sufficiency part is essentially Blyth's method, but the restriction to a fixed set B and to priors with compact support play a key role in what follows. The condition that $G_n(B) \ge 1$ ensures that the prior mass does not escape; note that the normalization of G_n does not affect δ_{G_n}, compare (2). Thus admissibility requires an approximation by Bayes rules in a sense that is sufficiently strong that even with $G_n(B) \ge 1$, the integrated risk difference (5) approaches zero.

A first consequence of this characterization is that in the present setting, any admissible rule δ is necessarily generalized Bayes: there exists a measure $F(d\theta)$ such that

$$\delta(x) = \delta_F(x) = x + \frac{\nabla f^*}{f^*}(x) \tag{6}$$

for all x. Brown establishes a continuity theorem for multivariate Laplace transforms and applies it to measures derived from G_n to show the existence of F, extending a univariate result of Sacks (1963). The result was extended to exponential families by Berger and Srinivasan (1978).

Brown's identity. The question now is which estimators δ_F are admissible. Substituting the Tweedie representations (6) and (3) into regret formula (4) and defining $h^* = g^*/f^*$ and $\hat{j} = \sqrt{h^*}$, we obtain

$$\begin{aligned} B(G,\delta_F) - B(G,\delta_G) &= \int |\delta_F - \delta_G|^2 g^* \\ &= \int \frac{|\nabla h^*|^2}{h^*} f^* \\ &= 4\int |\nabla \hat{j}|^2 f^*. \end{aligned} \tag{7}$$

This is the fundamental identity: it relates approximation of a fixed estimator δ_F in Bayes regret to the minimization of an "energy integral" over suitable functions $\hat{j}$, with the weight function f^* remaining fixed. [2]

For an admissible estimator $\delta = \delta_F$, let $\hat{j}_n$ denote the functions derived in this way from a sequence of priors G_n satisfying (5). Certain properties of $\hat{j}_n$ can be deduced from those of the priors G_n. For convenience here, assume that $\operatorname{supp}(F) = \mathbb{R}^p$. Since G_n has compact support, it follows from the smoothing effect of convolution that

$$\hat{j}_n^2(x) = \frac{g_n^*}{f^*}(x) = \frac{G_n \star \phi}{F \star \phi}(x) \to 0,$$

as $|x| \to \infty$. We write $j_\infty = \lim_{|x|\to\infty} j(x)$ below. The condition $G_n(B) \ge 1$ entails, perhaps after a one-time renormalization of $F(d\theta)$, that

$$\hat{j}_n(x) \ge 1, \quad |x| \le 1.$$

Thus each $\hat{j}_n$ belongs to the class of functions

$$\mathcal{J} := \{j(x) : j|_{\partial B} \ge 1, j_\infty = 0\}.$$

Brown's identity and these remarks show that

$$\delta_F \text{ admissible} \quad \Rightarrow \quad \inf_{j\in\mathcal{J}} \int |\nabla j|^2 f^* = 0.$$

[1]Written as $\delta_G(x) = x + \nabla(\log g^*)(x)$, it is credited by Robbins (1956), in a form applying to general exponential families, to Tweedie (1947). Efron (2011) dubs this "Tweedie's formula" and explores its use for correcting selection bias.

[2]Bickel (1981) is an example of the later use of this identity, in this case to the bounded normal mean problem.

Euler–Lagrange equation. The previous variational problem is linked to a partial differential equation by a standard result in the calculus of variations. Let D be a bounded open set with boundary ∂D. If k_D minimizes

$$\int_D |\nabla k|^2 f^* \quad \text{s.t. } k = k_0 \text{ on } \partial D,$$

then k_D is the *unique* solution to the boundary value problem

$$Lk = 0, \quad x \in D; \quad k = k_0 \quad \text{on } \partial D, \tag{8}$$

where the (elliptic, second-order) operator

$$Lk = \Delta k + \frac{\nabla f^*}{f^*} \cdot \nabla k. \tag{9}$$

Of particular interest are the annuli $D_n = \{1 < |x| < n\}$ with boundary conditions $k_0 = 1$ and 0 for $|x| = 1$ and n, respectively. We write $k_n(x)$ for the solution to (8) in this case. For example, if $f^* = 1$, corresponding to $\delta_F(x) = x$, then

$$k_n(x) = \begin{cases} 1 - \dfrac{\log|x|}{\log n} & p = 2, \\ \dfrac{|x|^{2-p} - n^{2-p}}{1 - n^{2-p}} & p \geq 3. \end{cases} \tag{10}$$

Brownian diffusion. The connection with recurrence is made by defining a continuous path process $\{Z_t, t \geq 0\}$ built from Brownian motion B_t via a stochastic differential equation:

$$dZ_t = \sqrt{2}\, dB_t + \frac{\nabla f^*}{f^*}(Z_t)\, dt, \quad Z_0 = x.$$

This process has local covariance $2I_p$ and local mean $(\nabla f^*/f^*)(x) = \delta_F(x) - x$. Thus, the Brownian motion is locally modified by a drift corresponding to the departure from the MLE. Johnstone and Lalley (1984) called this a *Brownian diffusion* in honor of both Browns. The elliptic operator (8), (9) appears as the infinitesimal generator

$$\lim_{t \to 0} \frac{\mathbb{E}^x k(Z_t) - k(x)}{t} = Lk(x).$$

If the process starts at $x \notin B$, it returns at some time[3] to $B = \{|x| \leq 1\}$ with *return probability* $\mathcal{K}(x) = \mathbb{P}^x\{\inf_{t \geq 0} |Z_t| = 1\}$. The diffusion is called *recurrent* if $\mathcal{K}(x) \equiv 1$ and *transient* otherwise.

Characterizing recurrence. Dynkin's formula is a stochastic process analog of the fundamental theorem of calculus: for a smooth function k and stopping time τ with $\mathbb{E}^x \tau < \infty$, it reads

$$\mathbb{E}^x k(Z_\tau) = k(x) + \mathbb{E}^x \int_0^\tau Lk(Z_s)\, ds.$$

Consider again the annulus D_n, and the first exit time τ_n, the first hitting time of the boundary ∂D_n. Recall that k_n minimizes $\int_{D_n} |\nabla k|^2 f^*$ with boundary conditions $k_n = 1$ for $|x| = 1$ and $k_n = 0$ for $|x| = n$. Since k_n also satisfies (8), Dynkin's formula shows that $k_n(x)$ increases as $n \to \infty$ to the recurrence probability:

$$\begin{aligned} k_n(x) &= \mathbb{E}^x k_n(Z_{\tau_n}) \\ &= \mathbb{P}^x\{Z_t \text{ hits } |x| = 1 \text{ before } |x| = n\} \\ &\nearrow \mathbb{P}^x\Big\{\inf_{t \geq 0} |Z_t| = 1\Big\} = \mathcal{K}(x). \end{aligned}$$

See Figure 1. This suggests the main result proved in Brown (1971), Section 4: that

$$\{Z_t\} \text{ is recurrent} \quad \text{if and only if}$$
$$\inf_{j \in \mathcal{J}} \int |\nabla j|^2 f^* = 0.$$

This general notion of a probabilistic solution of partial differential equations didn't originate with Brown, but he saw that it applied in the statistically derived setting, and needed to prove that it worked for his case, and that it provided an exact characterization of recurrence in terms of the variational problem.

Proving admissibility. To close the circle, it remains to show that recurrence of $\{Z_t\}$ implies admissibility. In this case, as just seen, $\int |\nabla k_n|^2 f^* \to 0$. Thus, it is natural to propose the priors

$$G_n(d\theta) = k_n^2(\theta) F(d\theta), \tag{11}$$

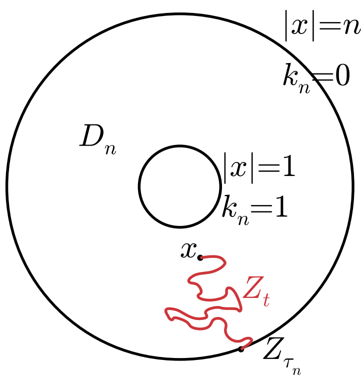

FIG. 1. *Dirichlet problem for an annulus D: $k_n(x)$ gives the probability that diffusion Z_t started at x exits D for the first time at the inner boundary.*

[3] If the process "explodes" in finite time $\tau_\infty < \infty$, it is declared to stay at ∞ thereafter: this (unusual) case is discussed in the correction note Brown (1973).

and to set

$$\hat{j}_n^2(x) = \frac{G_n \star \phi}{F \star \phi}(x) = \mathbb{E}_F[k_n^2(\theta)|x] \approx k_n^2(x)$$

at least heuristically, since $\mathbb{E}_F[\theta|x] = x + (\nabla f^*/f^*)(x)$. Recalling the criterion (5) and fundamental equation (7), to prove admissibility it would suffice, for example, to show that for some constant C,

$$B(\delta_{G_n}, \delta_F) - B(\delta_{G_n}, \delta_{G_n}) = 4 \int |\nabla \hat{j}_n|^2 f^* \le C \int |\nabla k_n|^2 f^* \to 0.$$

In his *Stat. Sci.* interview (DasGupta, 2005), Brown recalls that he was stuck at this point until a conversation with his Cornell applied mathematics colleague Jim Bramble. The Gauss mean value property for harmonic functions says that on any ball $B(x,r) = \{y : |y-x| < r\}$ interior to where $\Delta j(y) = 0$, it holds that $j(x) = \int_{B(x,r)} j(y)\, dy/\operatorname{vol}(B)$. Bramble's (unpublished) extension of this result was just what Brown needed. Under the assumption that $|\nabla f^*/f^*(x)| \le B$, it showed that there exist probability densities $r_x(y) \le C_{p,B}$ supported in $B(x, \frac{1}{2})$ so that

$$Lj(x) = 0 \quad \Rightarrow \quad j(x) = \int j(y) r_x(y)\, dy.$$

The key part of this result is that the density $r_x(y)$ satisfies a uniform bound. With it, and extra work, Brown arrived at the main technical result, Brown (1971), Theorem 5.1.1: if $|\nabla f^*/f^*(x)| \le B$, then

$$\begin{aligned} &Z_t \text{ recurrent} \\ &\quad \Rightarrow \quad \inf_{\mathcal{J}} \int |\nabla j|^2 f^* = 0 \\ &\quad \Rightarrow \quad \delta_F(x) = x + \frac{\nabla f^*}{f^*}(x) \text{ is admissible.} \end{aligned}$$

The final Section 6 of Brown (1971) explores a number of interesting and influential consequences for admissibility. In the one-dimensional case, the Euler–Lagrange equation can be explicitly solved and the general Theorem 5.1.1 can be reduced to an integral test. Thus for $\delta = \delta_F$, if either

$$\int_1^\infty 1/f^*(x)\, dx < \infty \quad \text{or} \quad \int_{-\infty}^{-1} 1/f^*(x)\, dx < \infty,$$

then δ_F is inadmissible. Conversely, if also $(d/dx)(\log f^*)$ is bounded, then divergence of both integrals above implies admissibility. Brown (1979a) gave a (necessarily) weird counterexample, using a discrete prior with an infinite number of widely spaced atoms, to show that the boundedness condition could not be entirely removed.

Turning to higher dimensions, the results of Section 6 have provided the basis for a key heuristic guiding modern shrinkage methodology using hierarchical Bayes priors:

- Proper priors yield admissible estimators.
- Too diffuse improper priors yield inadmissible estimators.
- Priors 'on the boundary of admissibility' are typically exactly balanced between being too vague and too concentrated.

To illustrate: given a general prior $F(d\theta)$ on $\mathbb{R}^m$, which might come from an hierarchical model, use the marginal density $f^*(x)$ to construct scalar averages

$$\bar{f}(r) = \int f^*(x) \mu_r(dx),$$

$$\underline{f}(r) = \int [f^*(x)]^{-1} \mu_r(dx)$$

over the surface of the sphere of radius r in $\mathbb{R}^m$. Combining Theorem 6.4.3 and (a special case of) Theorem 6.4.4, we can say that (a) δ_F is admissible if $\delta_F(x) - x$ is uniformly bounded and

$$\int_1^\infty r^{1-m} [\bar{f}(r)]^{-1}\, dr = \infty,$$

and (b) δ_F is inadmissible if

$$\int_1^\infty r^{1-m} \underline{f}(r)\, dr < \infty.$$

These criteria have been used in at least Berger and Strawderman (1996) and Berger, Strawderman and Tang (2005) to give concrete illustrations of the heuristic.

Brown's paper of course inspired some subsequent work on this and related problems. Srinivasan (1981) worked with Brown's boundary value problem characterization and showed that it held under notably weaker conditions on the risk function of δ_F, while Srinivasan (1982) used related ideas to pass between admissibility in estimating a natural parameter in exponential families and the multinormal mean problem. Johnstone (1984, 1986) considered analogs of the characterization in the discrete setting of Poisson observations. Eaton (1992) obtained a relationship between admissibility and recurrence of associated Markov chains in the context of "quadratically regular" decision problems. Hartigan (2012) evaluated priors using a Kullback–Leibler loss function and found an asymptotic characterization of admissibility in terms of solvability of associated elliptic differential equations.

Finally, we give an example of a later statistical result whose proof relied entirely on the recurrence characterization. Let $\delta(x)$ be an admissible estimator for $X \sim N_p(\theta, I)$. We can say, following Gutmann (1984), that δ is *immune* to the Stein effect if for any admissible estimator $\delta'(y)$ based on an independent observation $Y \sim N_q(\theta', I)$, the combined estimator (δ, δ') is admissible in the combined problem of estimating (θ, θ') under added squared error loss. Clearly, the MLE in dimensions 1 or 2 is not immune. Brown (1971), Section 6.5, showed that any proper Bayes estimator was immune, and gave two proofs, one statistical, and one based on recurrence. He left open the question of whether there existed *any* δ_F which is not proper Bayes and which is immune. The question was answered in the negative by Johnstone and Lalley (1984), and Brown's recurrence characterization was the foundation for the proof.

We note here also that Brown and Hwang (1986) showed that Stein's phenomenon could not occur (coordinatewise admissible estimators are immune) if, along with further regularity conditions on the decision problem, the parameter space is compact. The proof is direct, in that it does not use the recurrence characterization.

3. CLASSES OF LOSS FUNCTIONS

In a couple of papers, Brown considered admissibility relative to a class of loss functions. Suppose that $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, \dots, p$ are independent, with the variances σ_i^2 known. A rule is weakly admissible, or $\mathcal{L}$-admissible with respect to a class $\mathcal{L}$ if $R_L(\mu, \delta) \le R_L(\mu, \delta')$ for all $\mu \in \mathbb{R}^p$ and $L \in \mathcal{L}$ implies $R_L(\mu, \delta) \equiv R_L(\mu, \delta')$ for all $\mu \in \mathbb{R}^p$, $L \in \mathcal{L}$. Brown (1975) considers

$$\mathcal{L}(C) = \left\{ L(\mu, \delta) = \sum_{i=1}^p c_i(\mu_i - \delta_i)^2, \right.$$
$$\left. \text{for } c = (c_1, \dots, c_p) \in C \right\},$$

aiming to model a situation in which the relative values of c_i linking the independent problems cannot be specified with certainty. He gives a new proof of Stein's inadmissibility result for $\delta_0(x) = x$ (in the case C is a single point) using a notion of "tail domination" which is extended here to obtain the main result—namely necessary and sufficient conditions on C for δ_0 to be $\mathcal{L}(C)$ admissible. Some important examples: if C is the whole nonnegative quadrant ("total incompatibility"), then δ_0 is admissible, while if $c_i/c_j \le K$ for all i, j and finite K ("partial incompatibility"), then δ_0 is inadmissible.

Much later, Brown and Hwang (1989) considered the rather different class

$$\mathcal{L} = \{ L((\delta - \theta)' Q(\delta - \theta)) :$$
$$L(\cdot) \text{ any nondecreasing function} \}$$

for a now fixed positive definite matrix Q, and $\sigma_i^2 \equiv 1$. The surprise here is that δ_0 is $\mathcal{L}$-admissible *for any* p when $Q = I$, but for $Q \ne I$, it is inadmissible for p sufficiently large.

4. A UNIFIED ADMISSIBILITY CRITERION

Brown and Hwang (1982) gives a unified and remarkably simple yet general criterion for admissibility of generalized Bayes estimators for the mean vector in exponential families under a quadratic form loss. We will state the special case that applies to the multivariate normal mean setting with $X \sim N_p(\theta, I)$ in order to draw the connection with Brown (1971). Berger and DasGupta (2019), Section 5, in this volume has further discussion. Suppose that the prior $F(d\theta) = f(\theta)\,d\theta$ satisfies the growth condition

$$\int_{|\theta|>1} \frac{f(\theta)}{|\theta|^2 \log^2(|\theta| \vee 2)}\,d\theta < \infty, \tag{12}$$

where $a \vee b = \max(a, b)$, and a flatness condition

$$\int \frac{|\nabla f(\theta)|^2}{f(\theta)}\,d\theta < \infty. \tag{13}$$

If the generalized Bayes estimator δ_F also has bounded risk on compact sets, then it is admissible for squared error loss.

The proof is built upon Blyth's method. The choice of priors can be seen to be inspired by Brown (1971). Indeed, here $G_n(d\theta) = \check{k}_n^2(\theta)F(d\theta)$ is inspired by the choice (11), but instead of using the solutions of (8), we now use the simpler solution (10) to the *two-dimensional* Laplace equation:

$$\check{k}_n(\theta) = 1 - \frac{\log|\theta|}{\log n}, \quad 1 < |\theta| < n$$

and equal to 1 or 0 as $|\theta| \le 1$ or $\ge n$.

This simpler choice turns out to work perfectly well under the integrability conditions (12)–(13). It yields the quickest proof of admissibility of $\delta(x) = x$ for $p = 2$, and also for priors with $g(\theta) \le |\theta|^{2-p}$ in dimensions $p \ge 3$ under reasonable extra conditions.

5. DIFFERENTIAL INEQUALITIES

Brown (1979b) is a major paper that starts with

> Two interesting general features stand out in results so far obtained. One is that the admissibility of a generalized Bayes estimator seems to depend on the general structure of the problem (the general type of problem-location, scale, etc.—and the dimension of the parameter space) and on the generalized Bayes prior, but not on other more specific features of the problem such as the loss function used or the exact shape of the densities (normal, exponential, etc.). A similar comment holds for best invariant estimators in problems where they exist. A more striking feature is that certain estimators which have in the past seemed intuitively reasonable have turned out to be inadmissible.

Brown shows that the admissibility problem is related to a differential inequality. The qualitative relations among the coefficients appear to capture the general features determining admissibility, and are not affected by the more specific features.

The discussion proceeds at a heuristic level, with many of the (necessary) approximations rendered reasonable by detailed experience in earlier papers, and also with a number of particular applications described later in the paper: namely estimation of:

- several location parameters, with and without nuisance parameters,
- several Poisson means,
- the largest of several ordered translation paramters,
- a normal variance when the mean is unknown.

To give a flavor, suppose that the scalar random variable X has density $p_\xi(x)$ and an estimator $\delta(x) = x + \gamma(x)$ of ξ is evaluated using loss function $W(\delta(x) - \xi)$. [Brown also considers vector observations and parameters, and allows for presence of nuisance parameters.] Taylor expansions show that the risk difference

$$\Delta = \mathbb{E}\big[W(x + \gamma - \xi) - W(x + \gamma + \lambda - \xi)\big]$$
$$\doteq \mathbf{Q}\lambda - \frac{1}{2} n \lambda^2,$$

where the (linear) differential operator

$$\mathbf{Q}\lambda = -(m_1 + \psi)\lambda - m_2\, d\lambda/d\xi$$

is defined in terms of the moments $m_i = \mathbb{E}_\xi (x - \xi)^{i-1} W(x - \xi)$ while $\psi = \mathbb{E}_\xi \gamma(x) W''(x - \xi)$, and $n = \mathbb{E}_\xi W''(x - \xi)$. [When nuisance parameters are present, $\mathbf{Q}$ also includes second-order derivatives.] To show that δ is inadmissible, one may look first for a solution to the inequality $\mathbf{Q}\lambda - n\lambda^2/2 \geq 0$ as a prelude to a rigorous verification.

A second, related method is introduced in cases where δ is generalized Bayes for a prior, say with density g. It involves the adjoint operator

$$\mathbf{Q}^* g = -(m_1 + \psi) g + \frac{d}{d\xi}(m_2 g).$$

An important tool is a admissibility alternative formally proved in Brown (1980b), to the effect that if the risk difference Δ between δ_0 and δ' is asymptotically positive in the sense that $H(\xi)\Delta(\xi) \geq c > 0$ for $|\xi|$ large, then either δ_0 is inadmissible, or it is generalized Bayes for a prior $G(d\xi)$ satisfying $\int H^{-1}(\xi) G(d\xi) < \infty$. The strategy for proving inadmissibility for a generalized Bayes rule is then to show that the second case cannot occur.

Stein's Unbiased Estimate of Risk says that the mean squared error of an estimator $\delta(x) = x + \gamma(x)$ when $X \sim N_p(\theta, I)$ is given by

$$R(\theta, \delta) = p + \mathbb{E}\{R_0 \gamma(X)\},$$
$$R_0 \gamma(x) = 2\nabla \cdot \gamma(x) + |\gamma(x)|^2.$$

With $\gamma(x) = -(p-2)|x|^{-2}$, a short calculation yields $R_0\gamma < 0$ and so an immediate proof of the inadmissibility of $\delta_0(x) = x$.

Although published in the then Eastern bloc in Stein (1974), the identity was better known in the West through Stein (1981); indeed there is no mention of the result in Brown (1979b). Brown's Purdue IV paper Brown (1988a), however, responds with a detailed study of the unbiased risk estimator technique, and the relation between differential inequalities and admissibility.

Brown considers an initial estimator $\delta_\gamma = x + \gamma(x)$ and a candidate improvement $\delta_{\gamma+\lambda} = x + \gamma(x) + \lambda(x)$. The unbiased estimator for the difference in risks

$$R(\theta, \delta_{\gamma+\lambda}) - R(\theta, \delta_\gamma) = \mathbb{E}\{R_\gamma \lambda(X)\}$$

leads to consideration of the operator

$$R_\gamma \lambda = 2\nabla \cdot \lambda + 2\gamma \cdot \lambda + |\lambda|^2.$$

Certainly, if a solution λ of the inequality $R_\gamma\lambda \leq 0$ exists, then $\delta_{\gamma+\lambda}$ is at least as good as δ_γ and so, via a convexity argument, δ_γ is inadmissible. Brown's paper is concerned with the converse: he shows that if there is no λ making $R_\gamma\lambda \leq 0$, then admissibility of δ_γ is a reasonable conjecture, which might then be rigorously verified by other techniques.

Differential operators such as $R_0\gamma$ are treated as if they were actually the risk of an estimator corresponding to γ. Theorems 5.1 and 5.3 provide necessary and sufficient conditions for admissibility in this sense. The former shows that admissible rules are generalized Bayes with an analytic representation such as (6). The latter has the flavor of the variational characterization of admissibility in Brown (1971), now in the differential inequality setting.

One sample consequence (Example 9) is that no estimator improving on the James–Stein positive part rule $\delta_\gamma(x) = (1 - (p-2)|x|^{-2})_+ x$ can be found by the unbiased risk estimator technique. Another (Example 10) considers a Brownian motion $W(t)$ with mean μt and variance t. The uniform prior Bayes estimator $W(t)/t$ for μ under squared error loss is admissible using a fixed sample size $t = m$ say, but surprisingly is inadmissible if one stops at the first crossing of the SPRT boundaries or at $t = m$, whichever is earlier. This result is striking to objective Bayesians who would routinely use the uniform Bayes prior in this problem. It adds to the list of Brown's inadmissibility examples that continue to pose foundational challenges; see Berger and DasGupta (2019) in this volume for this, and further perspective on Brown (1979b) and Brown (1988a).

6. COMPLETE CLASS THEOREMS AND SEQUENTIAL PROBLEMS

Complete class theorems for a given statistical decision problem aim to give concrete descriptions of classes of decision rules within which all admissible procedures must lie. One example is the generalized Bayes representation for all admissible rules in the multivariate normal mean setting of Brown (1971); cf. (6) above. Brown published more than ten papers on complete class theorems; only the briefest categorization is attempted here (see also Berger and DasGupta, 2019):

- decision problems for density families with monotone likelihood ratio, Brown, Cohen and Strawderman (1976), Brown and Cohen (1995),
- estimation problems with finite sample spaces. The influential paper (Brown, 1981) gave a stepwise representation for admissible estimators in terms of proper Bayes rules. Brown, Chow and Fong (1992) considered the particular case of estimating a binomial variance, while Ighodaro, Santner and Brown (1982) studied multinomial estimation under two natural loss functions,
- estimation problems with countable discrete sample spaces. Brown and Farrell (1985a) and Brown and Farrell (1988) gave a stepwise representation for admissible estimators, this time in terms of generalized Bayes rules. A related paper (Brown and Farrell, 1985b) determines which affine estimators $Mx + \gamma$ are limits of Bayes estimators and which are admissible,
- hypothesis testing problems with simple null hypothesis, Brown and Marden (1989) and Brown and Marden (1992),
- sequential testing problems, to which we now turn.

Brown wrote a number of papers on sequential problems; the summary here is again brief. Brown (1977) is concerned with the general decision theory of problems in which the observations are taken sequentially. The other papers in this group consider sequential tests, where typically the risk function is either a linear combination of the probability of error, or a bivariate vector with these two terms as separate components. Brown, Cohen and Strawderman (1979a) shows that admissibility of fixed sample size tests, when considered in a sequential setting, depends on which of the risk functions is used.

Complete class theorems are given in Brown, Cohen and Strawderman (1980) and Berk, Brown and Cohen (1981). Monotonicity properties of sequential tests are studied for one-sided hypotheses in Brown, Cohen and Strawderman (1979b), and for a two-sided hypothesis in Brown and Greenshtein (1992). For the sequential one-sided testing problem, Brown and Cohen (1981) shows inadmissibility of tests with unbounded continuation region, while Brown, Cohen and Samuel-Cahn (1983) characterizes which sequential probability ratio tests (SPRTs) are admissible.

7. CONCLUDING REMARKS

We mention here briefly some other papers by Brown relating to admissibility.

Brown and Cohen (1974) considered the problem of estimating the common mean of two independent normal distributions, each with unknown variances, on the basis of samples of size m and n. They showed that the sample mean of the first population could be improved on (with an unbiased estimator having smaller variance), provided $m \geq 2$ and $n \geq 3$, but not if $n = 2$. The then popular problem of "recovery of interblock information" in balanced incomplete block designs can be viewed as an example of the common mean problem.

Brown (1988b) is one of the first papers to consider admissibility in the context of finite sample size invariant nonparametric estimation of a distribution function and of its median. Many questions are resolved for Cramér–von Mises and Kolmogorov–Smirnov loss functions, while others are left open.

Brown's important Wald Lectures article (Brown, 1990) on the inadmissibility of the least squares estimate of the intercept in random design regression is discussed in Berger and DasGupta (2019), Section 5.

Brown et al. (2006) presents two expectation identities, one of which uses the heat equation and is equivalent to Stein's identity, Stein (1981). A cornucopia of applications are given, including inadmissibility results and a Stein *inequality* for spherically symmetric t-distributions.

Brown, George and Xu (2008) studies admissibility of predictive density estimators: given independent p-dimensional normal vectors X and Y with common unknown mean, one seeks to estimate the predictive density of Y given X using Kullback–Leibler loss. The substantial parallels between this problem and the multivariate normal mean problem for squared error loss (first noted in George, Liang and Xu (2006)) are extended by establishing complete class theorems, and a remarkable analog of the sufficient conditions for admissibility of Brown and Hwang (1982).

Among the topics not covered here [see Berger and DasGupta, 2019 for the first two]:

- discussions of admissibility and complete classes contained in Brown's monograph on exponential families and unpublished notes on decision theory,
- Brown's papers on minimax estimation that don't directly address admissibility,
- Brown's more recent work on SURE estimates for heteroscedastic models, Brown 2008; Xie–Kou–Brown 2012, 2016; Weinstein–Ma–Brown–Zhang 2017.

ACKNOWLEDGMENTS

Supported in part by NSF Grant DMS-1407813 and 1811614. In preparing this review, I have benefitted from the extensive discussion of admissibility in Lehmann and Casella (1998) and Berger (1985), as well as from the careful summaries of some individual papers in *Mathematical Reviews* by reviewers such as Steven Arnold, Jim Berger, Arthur Cohen, Anirban DasGupta, Roger Farrell, Kurt Hoffmann, Andrew Rukhin, Jerry Sacks and Lionel Weiss. A referee also provided very helpful comments incorporated in this revision.

REFERENCES

BERGER, J. (1980). Improving on inadmissible estimators in continuous exponential families with applications to simultaneous estimation of gamma scale parameters. *Ann. Statist.* **8** 545–571. MR0568720

BERGER, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR0804611

BERGER, J. and DASGUPTA, A. (2019). Larry Brown's contributions to parametric inference, decision theory and foundations: A survey. *Statist. Sci.* **34** 621–634.

BERGER, J. O. and SRINIVASAN, C. (1978). Generalized Bayes estimators in multivariate problems. *Ann. Statist.* **6** 783–801. MR0478426

BERGER, J. O. and STRAWDERMAN, W. E. (1996). Choice of hierarchical priors: Admissibility in estimation of normal means. *Ann. Statist.* **24** 931–951. MR1401831

BERGER, J. O., STRAWDERMAN, W. and TANG, D. (2005). Posterior propriety and admissibility of hyperpriors in normal hierarchical models. *Ann. Statist.* **33** 606–646. MR2163154

BERK, R. H., BROWN, L. D. and COHEN, A. (1981). Properties of Bayes sequential tests. *Ann. Statist.* **9** 678–682. MR0615445

BICKEL, P. J. (1981). Minimax estimation of the mean of a normal distribution when the parameter space is restricted. *Ann. Statist.* **9** 1301–1309. MR0630112

BLACKWELL, D. (1951). On the translation parameter problem for discrete variables. *Ann. Math. Stat.* **22** 393–399. MR0043418

BLYTH, C. R. (1951). On minimax statistical decision procedures and their admissibility. *Ann. Math. Stat.* **22** 22–42. MR0039966

BROWN, L. D. (1966). On the admissibility of invariant estimators of one or more location parameters. *Ann. Math. Stat.* **37** 1087–1136. MR0216647

BROWN, L. (1968). Inadmissiblity of the usual estimators of scale parameters in problems with unknown location and scale parameters. *Ann. Math. Stat.* **39** 29–48. MR0222992

BROWN, L. D. (1971). Admissible estimators, recurrent diffusions, and insoluble boundary value problems. *Ann. Math. Stat.* **42** 855–903. MR0286209

BROWN, L. D. (1973). Correction to: "Admissible estimators, recurrent diffusions, and insoluble boundary value problems" (*Ann. Math. Statist.* **42** (1971), 855–903). *Ann. Statist.* **1** 594.

BROWN, L. D. (1975). Estimation with incompletely specified loss functions (the case of several location parameters). *J. Amer. Statist. Assoc.* **70** 417–427. MR0373082

BROWN, L. D. (1977). Closure theorems for sequential-design processes. In *Statistical Decision Theory and Related Topics, II* (*Proc. Sympos., Purdue Univ., Lafayette, Ind.*, 1976) 57–91. MR0461738

BROWN, L. D. (1979a). Counterexample—an inadmissible estimator which is generalized Bayes for a prior with "light" tails. *J. Multivariate Anal.* **9** 332–336. MR0538413

BROWN, L. D. (1979b). A heuristic method for determining admissibility of estimators—with applications. *Ann. Statist.* **7** 960–994. MR0536501

BROWN, L. (1980a). Examples of Berger's phenomenon in the estimation of independent normal means. *Ann. Statist.* **8** 572–585. MR0568721

BROWN, L. D. (1980b). A necessary condition for admissibility. *Ann. Statist.* **8** 540–544. MR0568719

BROWN, L. D. (1981). A complete class theorem for statistical problems with finite sample spaces. *Ann. Statist.* **9** 1289–1300. MR0630111

BROWN, L. D. (1988a). The differential inequality of a statistical estimation problem. In *Statistical Decision Theory and Related Topics, IV, Vol.* 1 (*West Lafayette, Ind.*, 1986) 299–324. Springer, New York. MR0927109

BROWN, L. D. (1988b). Admissibility in discrete and continuous invariant nonparametric estimation problems and in their multinomial analogs. *Ann. Statist.* **16** 1567–1593. MR0964939

BROWN, L. D. (1990). The 1985 Wald Memorial Lectures. An ancillarity paradox which appears in multiple linear regression. *Ann. Statist.* **18** 471–538. MR1056325

BROWN, L. D., CHOW, M. S. and FONG, D. K. H. (1992). On the admissibility of the maximum-likelihood estimator of the binomial variance. *Canad. J. Statist.* **20** 353–358. MR1208348

BROWN, L. D. and COHEN, A. (1974). Point and confidence estimation of a common mean and recovery of interblock information. *Ann. Statist.* **2** 963–976. MR0356334

BROWN, L. D. and COHEN, A. (1981). Inadmissibility of large classes of sequential tests. *Ann. Statist.* **9** 1239–1247. MR0630106

BROWN, L. D. and COHEN, A. (1995). Complete classes for confidence set estimation. *Statist. Sinica* **5** 291–301. MR1329299

BROWN, L. D., COHEN, A. and SAMUEL-CAHN, E. (1983). A sharp necessary condition for admissibility of sequential tests—necessary and sufficient conditions for admissibility of SPRTs. *Ann. Statist.* **11** 640–653. MR0696075

BROWN, L. D., COHEN, A. and STRAWDERMAN, W. E. (1976). A complete class theorem for strict monotone likelihood ratio with applications. *Ann. Statist.* **4** 712–722. MR0415821

BROWN, L. D., COHEN, A. and STRAWDERMAN, W. E. (1979a). On the admissibility or inadmissibility of fixed sample size tests in a sequential setting. *Ann. Statist.* **7** 569–578. MR0527492

BROWN, L. D., COHEN, A. and STRAWDERMAN, W. E. (1979b). Monotonicity of Bayes sequential tests. *Ann. Statist.* **7** 1222–1230. MR0550145

BROWN, L. D., COHEN, A. and STRAWDERMAN, W. E. (1980). Complete classes for sequential tests of hypotheses. *Ann. Statist.* **8** 377–398. Erratum **17** (1989), 1414–1416. MR0560735

BROWN, L. D. and FARRELL, R. H. (1985a). Complete class theorems for estimation of multivariate Poisson means and related problems. *Ann. Statist.* **13** 706–726. MR0790567

BROWN, L. D. and FARRELL, R. H. (1985b). All admissible linear estimators of a multivariate Poisson mean. *Ann. Statist.* **13** 282–294. MR0773168

BROWN, L. D. and FARRELL, R. H. (1988). Proof of a necessary and sufficient condition for admissibility in discrete multivariate problems. *J. Multivariate Anal.* **24** 46–52. MR0925128

BROWN, L. D. and FOX, M. (1974a). Admissibility of procedures in two-dimensional location parameter problems. *Ann. Statist.* **2** 248–266. MR0420929

BROWN, L. D. and FOX, M. (1974b). Admissibility in statistical problems involving a location or scale parameter. *Ann. Statist.* **2** 807–814. MR0370850

BROWN, L. D., GEORGE, E. I. and XU, X. (2008). Admissible predictive density estimation. *Ann. Statist.* **36** 1156–1170. MR2418653

BROWN, L. D. and GREENSHTEIN, E. (1992). Two-sided sequential tests. *Ann. Statist.* **20** 555–561. MR1150361

BROWN, L. D. and HWANG, J. T. (1982). A unified admissibility proof. In *Statistical Decision Theory and Related Topics, III, Vol.* 1 (*West Lafayette, Ind.*, 1981) 205–230. Academic Press, New York. MR0705290

BROWN, L. D. and HWANG, J. T. (1986). Some limitation on Stein's phenomenon. *Comm. Statist. Theory Methods* **15** 2025–2042. MR0851855

BROWN, L. D. and HWANG, J. T. (1989). Universal domination and stochastic domination: U-admissibility and U-inadmissibility of the least squares estimator. *Ann. Statist.* **17** 252–267. MR0981448

BROWN, L. D. and MARDEN, J. I. (1989). Complete class results for hypothesis testing problems with simple null hypotheses. *Ann. Statist.* **17** 209–235. MR0981446

BROWN, L. D. and MARDEN, J. I. (1992). Local admissibility and local unbiasedness in hypothesis testing problems. *Ann. Statist.* **20** 832–852. MR1165595

BROWN, L. D. and ZHAO, L. H. (2012). A geometrical explanation of Stein shrinkage. *Statist. Sci.* **27** 24–30. MR2953493

BROWN, L., DASGUPTA, A., HAFF, L. R. and STRAWDERMAN, W. E. (2006). The heat equation and Stein's identity: Connections, applications. *J. Statist. Plann. Inference* **136** 2254–2278. MR2235057

DASGUPTA, A. (2005). A conversation with Larry Brown. *Statist. Sci.* **20** 193–203. MR2183449

EATON, M. L. (1992). A statistical diptych: Admissible inferences—recurrence of symmetric Markov chains. *Ann. Statist.* **20** 1147–1179. MR1186245

EFRON, B. (2011). Tweedie's formula and selection bias. *J. Amer. Statist. Assoc.* **106** 1602–1614. MR2896860

FARRELL, R. H. (1968a). On a necessary and sufficient condition for admissibility of estimators when strictly convex loss is used. *Ann. Math. Stat.* **39** 23–28. MR0231484

FARRELL, R. H. (1968b). Towards a theory of generalized Bayes tests. *Ann. Math. Stat.* **39** 1–22. MR0224211

GEORGE, E. I., LIANG, F. and XU, X. (2006). Improved minimax predictive densities under Kullback–Leibler loss. *Ann. Statist.* **34** 78–91. MR2275235

GUTMANN, S. (1984). Decisions immune to Stein's effect. *Sankhyā Ser. A* **46** 186–194. MR0778869

HARTIGAN, J. A. (2012). Asymptotic admissibility of priors and elliptic differential equations. In *Contemporary Developments in Bayesian Analysis and Statistical Decision Theory: A Festschrift for William E. Strawderman. Inst. Math. Stat.* (*IMS*) *Collect.* **8** 117–130. IMS, Beachwood, OH. MR3202507

HODGES, J. L. JR. and LEHMANN, E. L. (1951). Some applications of the Cramér–Rao inequality. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 1950 13–22. Univ. California Press, Berkeley and Los Angeles. MR0044795

IGHODARO, A., SANTNER, T. and BROWN, L. (1982). Admissibility and complete class results for the multinomial estimation problem with entropy and squared error loss. *J. Multivariate Anal.* **12** 469–479. MR0680522

JAMES, W. and STEIN, C. (1961). Estimation with quadratic loss. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I* 361–379. Univ. California Press, Berkeley, CA. MR0133191

JOHNSTONE, I. (1984). Admissibility, difference equations and recurrence in estimating a Poisson mean. *Ann. Statist.* **12** 1173–1198. MR0760682

JOHNSTONE, I. (1986). Admissible estimation, Dirichlet principles and recurrence of birth–death chains on $\mathbf{Z}_+^p$. *Probab. Theory Related Fields* **71** 231–269. MR0816705

JOHNSTONE, I. and LALLEY, S. (1984). On independent statistical decision problems and products of diffusions. *Z. Wahrsch. Verw. Gebiete* **68** 29–47. MR0767442

KAKUTANI, S. (1944). On Brownian motions in n-space. *Proc. Imp. Acad.* (*Tokyo*) **20** 648–652. MR0014646

LEHMANN, E. L. and CASELLA, G. (1998). *Theory of Point Estimation*, 2nd ed. *Springer Texts in Statistics*. Springer, New York. MR1639875

LÉVY, P. (1940). Le mouvement brownien plan. *Amer. J. Math.* **62** 487–550. MR0002734

ROBBINS, H. (1956). An empirical Bayes approach to statistics. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 1954–1955, *Vol. I* 157–163. Univ. California Press, Berkeley and Los Angeles. MR0084919

RUKHIN, A. L. (1995). Admissibility: Survey of a concept in progress. *Int. Stat. Rev.* 95–115.

SACKS, J. (1963). Generalized Bayes solutions in estimation problems. *Ann. Math. Stat.* **34** 751–768. MR0150908

SRINIVASAN, C. (1981). Admissible generalized Bayes estimators and exterior boundary value problems. *Sankhya*, *Ser. A* **43** 1–25. MR0656265

SRINIVASAN, C. (1982). A relation between normal and exponential families through admissibility. *Sankhya*, *Ser. A* **44** 423–435. MR0705465

STEIN, C. (1955). A necessary and sufficient condition for admissibility. *Ann. Math. Stat.* **26** 518–522. MR0070929

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*, 1954–1955, *Vol. I* 197–206. Univ. California Press, Berkeley and Los Angeles. MR0084922

STEIN, C. (1959). The admissibility of Pitman's estimator of a single location parameter. *Ann. Math. Stat.* **30** 970–979. MR0109392

STEIN, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Ann. Inst. Statist. Math.* **16** 155–160. MR0171344

STEIN, C. (1974). Estimation of the mean of a multivariate normal distribution. In *Proceedings of the Prague Symposium on Asymptotic Statistics* (*Charles Univ.*, *Prague*, 1973), *Vol. II* 345–381. MR0381062

STEIN, C. M. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.* **9** 1135–1151. MR0630098

TWEEDIE, M. C. K. (1947). Functions of a statistical variate with given means, with special reference to Laplacian distributions. *Proc. Camb. Philos. Soc.* **43** 41–49. MR0018381