CHARLES STEIN AND INVARIANCE: BEGINNING WITH THE HUNT–STEIN THEOREM

BY MORRIS L. EATON[1] AND EDWARD I. GEORGE[2]

[1]*School of Statistics, University of Minnesota, eaton@stat.umn.edu*
[2]*Department of Statistics, University of Pennsylvania, edgeorge@wharton.upenn.edu*

When statistical decision theory was emerging as a promising new paradigm, Charles Stein was to play a major role in the development of minimax theory for invariant statistical problems. In some of his earliest work with Gil Hunt, he set out to prove that, in problems where invariant procedures have constant risk, any best invariant test would be minimax among all tests. Although finding it not quite true in general, this led to the legendary Hunt–Stein theorem, which established the result under restrictive conditions on the underlying group of transformations. In decision problems invariant under such suitable groups, an overall minimax test was guaranteed to reside within the class of invariant procedures where it would typically be much easier to find. But when it did not seem possible to establish this result for invariance under the full linear group, he instead turned to prove its impossibility with counterexamples such as the nonminimaxity of the usual sample covariance estimator where the full linear group was just too big for the Hunt–Stein theorem to apply. Further explorations of invariance such as the sometimes problematic inference under a fiducial distribution, or the characterization of a best invariant procedure as a formal Bayes procedure under a right Haar prior, are further examples of the far reaching influence of Stein's contributions to invariance theory.

1. Introduction. Invariance arguments in statistics date back at least to the work of Fisher and Hotelling, but perhaps the beginnings of a systematic development are the two highly original contributions of Pitman (1939a, 1939b). These two works treat estimation and testing problems for statistical models indexed by real valued translation and/or scale parameters. In that same year, in what was arguably the first paper on statistical decision theory, Wald (1939) introduced minimaxity and admissibility as desirable risk properties for the evaluation of estimation and testing procedures. He there considered Pitman's invariant procedure for the location problem, but was unsuccessful in his attempt to prove its minimaxity and admissibility. Four years later, while working on general decision theory problems, Wald (1943) introduced the related notion of stringency as an optimality property of statistical tests. Here, too, at least initially, it proved to be rather difficult to apply this notion to standard testing procedures. These developments, as well as personal encouragement from Wald himself, would set the stage for Charles Stein to focus on the landscape of risk properties of invariant statistical procedures, a focus that would persist throughout his career. Indeed, the Pitman work coupled with the challenge of establishing the stringency of testing procedures would lead to the initial Hunt–Stein work on invariant testing problems.

Brought together by chance in the middle of their graduate school studies to carry out weather forecasting for the U.S. Army during 1945–1946, Gilbert Hunt and Charles Stein collaborated to develop a general minimax theory for testing problems invariant under a group of transformations. Within this framework, they established what has come to be known as

the Hunt–Stein theorem and showed how to apply this discovery to the study of most stringent tests. Unfortunately, Hunt and Stein never published their results as their working manuscript got mislaid and lost, apparently while they were continuing to flesh out the details of some key counterexamples, (DeGroot (1986)). Fortunately, Stein communicated his results to Erich Lehmann, who discussed them and some of the applications in an early "theory of hypothesis testing" paper, which led to the wide dissemination of their ideas, (Lehmann (1950)). Perhaps the most striking thing about the Hunt–Stein result is that the main assumption involves the group of transformations and not the statistical model of the testing problem.

A natural outgrowth of the Hunt–Stein theorem was a concerted effort to establish a more comprehensive invariant minimax theorem. By this, we mean a result that implies that the minimax risk, M, of an invariant decision problem is equal to the invariant minimax risk, M_I, of the problem. The work of Peisakoff (1950), Kudo (1955) and Kiefer (1957) established various conditions under which $M = M_I$. The most limiting condition in these works is directly related to the Hunt–Stein condition on the group of transformations leaving the statistical problem invariant. This condition, which we denote by HS, is defined precisely in Section 2 below. In brief, when HS holds, then $M = M_I$. Of particular interest early on was whether this condition was met by Gl_p, the multiplicative group of all $p \times p$ nonsingular matrices, with p larger than one. Peisakoff (1950) asserted that HS holds for Gl_p but the proof contains a lacuna, Kudo (1955) said it was not known whether Gl_p satisfies HS, and then Kiefer (1957), p. 587, reported that Stein had constructed an example showing that Gl_p does not satisfy HS.

The Stein example alluded to by Kiefer was reported in an unpublished Stanford technical report issued in 1956. Indeed, this report, dealing almost exclusively with inferential problems of multivariate analysis, contains a variety of invariant examples where the invariant minimax theorem fails. The key example there focused on invariance under the group Gl_p when the estimation of a covariance matrix was of special interest. A related testing example of Stein's is also reported in Lehmann (1959), Example 9 on p. 338; also see problem 10 on p. 344. A significant portion of the material in Stein (1956a) is reported in James and Stein (1961).

An interesting aspect of the Pitman (1939a, 1939b) work is the interpretation of best invariant procedures in terms of a "fiducial" distribution. Related to this is an example in Stein (1959) where a fiducial calculation leads to an inference that is strongly at odds with a relevant frequentist calculation. This example leads to the conclusion that in certain situations, treating a natural fiducial distribution as a "regular" probability distribution can lead to inferential statements that are widely divergent from frequentist evaluations. However, a probability matching problem is treated in Stein (1965), pp. 221–225, where the calculation of a posterior distribution using a right Haar prior hints at more general representations of best invariant decision rules. In this context, the role of right Haar measure was mentioned in passing in Peisakoff (1950), pp. 37–38. Contemporaneous work of Hora (1964), reported in Hora and Buehler (1966), gives a more detailed account of some aspects of the Stein discussion although Stein (1965) is not referenced. The material in Stein (1965) was reported at a scientific meeting at Berkeley in 1963.

Charles Stein's work on invariance in the 1945–1965 period has been enormously influential and has inspired a wide range of theory and applications over the past 70 years. Unfortunately, the Hunt–Stein work and Stein (1956a) are unpublished, and his elegant "right Haar" argument in Stein (1965) is at best sketchy. In spite of this, the importance of these contributions is recognized by almost everyone interested in invariance. In the next three sections, we shall discuss the two unpublished works and the right Haar argument in a bit more detail. This work continues to influence some current research work on probability matching, incoherence and strong inconsistency, fiducial arguments and invariance theory in general.

2. The Hunt–Stein theorem and its generalizations. For invariant statistical testing problems as defined in Lehmann (1950), the central question Hunt and Stein sought to address, was whether a minimax invariant test would be minimax (i.e., maximize the minimum power) within the class of all tests of a given size. Initially setting out to prove that a best invariant test would be minimax in problems where invariant tests have constant risk, they quickly realized that such a property would hinge essentially on the theoretical properties of G, the group of transformations under consideration for the problem at hand, (DeGroot (1986)).

Although the original statement and proof of their results remain lost, Kiefer (1966) remarks that the following version of the Hunt–Stein theorem, reported by Lehmann (1959) in Lemma 2 on p. 332 and Theorem 3 on p. 336, is probably closest in spirit to the original.

THEOREM 2.1 (Hunt–Stein). *Let* $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ *be a dominated family of distributions on the sample space* $(\mathcal{X}, \mathcal{A})$, *and let* G *be a group of transformations of* $(\mathcal{X}, \mathcal{A})$ *such that the induced group* $\bar{G}$ *leaves the two subsets* Ω_H *and* Ω_K *of* Ω *invariant. Let* $\mathcal{B}$ *be a* σ-*field of subsets of* G *such that for any* $A \in \mathcal{A}$, *the set of pairs* (x, g) *with* $gx \in A$ *is in* $\mathcal{A} \times \mathcal{B}$, *and for any* $B \in \mathcal{B}$ *and* $g \in G$, *the set* Bg *is in* $\mathcal{B}$. *Suppose that there exists a sequence of probability distributions* ν_n *over* $(G, \mathcal{B})$, *which is asymptotically right invariant in the sense that for any* $g \in G$ *and* $B \in \mathcal{B}$,

$$(1) \qquad \lim_{n \to \infty} |\nu_n(Bg) - \nu_n(B)| = 0.$$

Then for any critical function ϕ, *there exists an* (*almost*) *invariant critical function* ψ, *which satisfies*

$$(2) \qquad \inf_{\bar{G}} E_{\bar{g}\theta}\phi(X) \le E_\theta \psi(X) \le \sup_{\bar{G}} E_{\bar{g}\theta}\phi(X)$$

for all $\theta \in \Omega$.

Just as Hunt and Stein had sought to establish, it follows directly from (2) that if there exists a level-α test ϕ_0 of $\theta \in \Omega_H$ versus $\theta \in \Omega_K$ that maximizes $\inf_{\Omega_k} E_\theta\phi(X)$, then there also exists an invariant test ψ with this property. (Note that as described in Lehmann (1959), Theorem 4 of Chapter 6, under suitable measurability assumptions which are "satisfied in all the usual applications," to every almost invariant procedure there is an equivalent invariant procedure.)

The essential property of G in the Hunt–Stein theorem is the existence of the asymptotically right invariant sequence ν_n in (1). We refer to this property as condition HS (for Hunt–Stein) following Bondar and Milnes (1981) who named it as such in their important survey paper. Condition HS is the essential ingredient needed to show that an invariant test ψ satisfying (2) can be obtained from ϕ_0 by averaging over the group G. This is transparent, for example, when $G = \{g_1, \dots, g_N\}$ is a finite group and $\nu_n \equiv \nu$ is the uniform distribution over G, which trivially satisfies (1). In this very simple case, the invariant critical function $\psi(x) = \frac{1}{N} \sum_1^N \phi_0(g_i x)$ is easily seen to satisfy (2) since $E_\theta\phi_0(gX) = E_{\bar{g}\theta}\phi_0(X)$ is an invariant average of terms, which includes both the upper and lower bounds of (2). To make such a construction of ψ work for larger groups, it is necessary to replace the invariant averaging operation $\frac{1}{N} \sum_{i=1}^N$ by a suitable mean preserving measure on G. For compact groups for which the invariant Haar measure has finite mass, this is straightforward. For noncompact groups, the condition HS or its equivalent is needed to asymptotically obtain suitable right invariant averaging.

For invariance under locally compact groups, statements of the Hunt–Stein theorem in the literature often substitute the property of amenability for condition HS. The notion of

amenability (existence of an invariant mean, as discussed in Greenleaf (1969)) was introduced by von Neumann in (1929). The paper of Bondar and Milnes (1981) contains a dozen or so conditions all of which are, modulo some minor regularity conditions, equivalent to amenability and HS. Essentially all of the research establishing these equivalents was done in the period 1946 to 1980 (see the eighty or so references in Bondar and Milnes).

A convenient sufficient condition for HS to hold is to let v_n be the restriction of a right Haar measure to a compact set K_n appropriately normalized so that v_n is a probability measure. For example, when G is the real line and $K_n = [-n, n]$, then HS holds when v_n is Lebesgue measure restricted to K_n and normalized by $(2n)$. As mentioned in Lehmann (1950), a useful condition that HS hold is the following. Assume N is a closed normal subgroup of G such that both N and the quotient group G/N both satisfy HS. Then G satisfies HS. Hence if G is a finite direct product of groups that satisfy HS, then so does G. This observation together with induction can be used to show, for example, that G_T, the group of $p \times p$ lower triangular matrices with positive diagonal elements, satisfies HS. In the hypothesis testing context, the first invariant statistical example where HS does not hold was given by Peisakoff (1950). The example involved the free group on two generators. Other such examples of more direct statistical interest were given in Stein (1956a). These are discussed in the next section.

For general invariant statistical decision problems as described, for example, in Stein (1956a) or Kiefer (1957), an immediate consequence of the Hunt–Stein work was interest in the validity of more comprehensive invariant minimax theorems. To state the issue precisely for an invariant decision problem, let $\mathcal{D}$ be the class of all decision rules and let $\mathcal{D}_I \subset \mathcal{D}$ be the subclass of all invariant decision rules. Given $\delta \in \mathcal{D}$, let $r(\delta, \theta)$ denote the risk of δ at θ. As usual, we define M and M_I by

$$(3) \qquad\qquad M = \inf_{\delta \in \mathcal{D}} \sup_{\theta} r(\theta, \delta),$$

$$(4) \qquad\qquad M_I = \inf_{\delta \in \mathcal{D}_I} \sup_{\theta} r(\theta, \delta).$$

These are the minimax risk and the invariant minimax risk, respectively. Obviously, $M \leq M_I$ and when $M = M_I$ we say that an invariant minimax theorem holds. In such cases, it follows that when δ is minimax within $\mathcal{D}_I$, it will also be minimax within $\mathcal{D}$. In a decision theoretic setting, depending on the structure of the particular problem at hand, finding a minimax $\delta \in \mathcal{D}_I$ is often much easier than finding a minimax $\delta \in \mathcal{D}$. Thus, the validity of an invariant minimax theorem provides a simplified route for finding both an overall minimax procedure and the minimax risk it attains. These at least serve as useful benchmarks against which competing procedures may be compared.

In essence, Hunt–Stein showed that $M = M_I$ for hypothesis testing problems when HS holds. Following on the heels of the Hunt–Stein contribution was the work of Peisakoff (1950), Kudo (1955) and Kiefer (1957) where versions of the invariant minimax theorem are given for various types of invariant decision problems. The three papers all have regularity conditions regarding the model, the loss function, the action of the group and the "size" of the group. What is striking is the similarity of the size assumptions with the HS condition. Indeed the size condition seems to be an invariant for these minimax theorems. Kiefer (1957) contains an excellent discussion of the different conditions and arguments, although Kudo (1955) is not discussed in detail. A general heuristic exposition of these developments can also be found in Kiefer (1966). Further discussion and illumination of the assumptions in Kiefer (1957) appear in Brown (1986), who also provides a sketch of an alternative proof of Kiefer's results using a fixed-point approach proposed by Huber. Further overviews and elaborations of Hunt–Stein theory appear in Berger (1985), Lehmann and Casella (1998) and Robert (2001).

3. Gl_p **is not amenable and related issues.**   Stein's work on invariant estimation yielded two fundamental results that were announced in Stein (1956a, 1956b). Perhaps the more famous is the inadmissibility of the maximum likelihood estimator (MLE) of the mean vector of a p-dimensional multivariate normal distribution when $p \geq 3$ (the loss is quadratic, and the MLE is also the best translation invariant estimator). Of course, this led to James and Stein (1961) and the James–Stein estimator. The other result deals with the estimation of a covariance matrix Σ based on a sample $X_1, X_2, \ldots, X_n$ from a p-dimensional mean zero multivariate normal distribution and $p \leq n$. For this estimation problem, Stein used the loss function

$$(5) \qquad L(\hat{\Sigma}, \Sigma) = \operatorname{tr}(\hat{\Sigma}\Sigma^{-1}) - \log \det(\hat{\Sigma}\Sigma^{-1}) - p.$$

With this loss, the problem of estimating Σ is invariant under both G_T and Gl_p. A sufficient statistic for this problem is $S = \sum_{i=1}^{n} X_i X_i'$. Because G_T is transitive on the parameter space of all $p \times p$ positive definite matrices, all G_T invariant estimators have constant risk. Of course, all Gl_p invariant estimators also have constant risk since G_T is a subgroup of Gl_p.

What Stein (1956a) showed was the following. First, the MLE $\frac{1}{n}S$ is the best Gl_p invariant estimator and has constant risk, say r_0. Next, write $S = TT'$ where T is the unique element of G_T satisfying this relationship. Then the minimum risk G_T invariant estimator of Σ is $\hat{\Sigma} = TDT'$ where D is a $p \times p$ diagonal matrix with diagonal elements $d_i = (n + p - 2i + 1)^{-1}$ for $i = 1, \ldots, p$. When $p \geq 2$, $\hat{\Sigma}$ differs from the MLE and is not invariant under Gl_p. The risk of this estimator is a constant, say r_1. The key result is that r_1 is strictly less than r_0. From the results of Kudo (1955), the group G_T is amenable so the minimax risk for this problem is r_1. Hence no Gl_p invariant estimator can be minimax, and thus Gl_p cannot be amenable.

The validity of similar Gl_p results for other models, problems and/or loss functions has been established. In the context of covariance estimation, see Selliah (1964), Olkin and Selliah (1977) and Eaton (1970). In addition, the de Finetti notion of incoherence (and its equivalent of Stone's strong inconsistency; see Eaton and Freedman (2004)) also arises in Gl_p invariant situations. For example, in classical prediction problems involving multivariate normal sampling, both Gl_p invariant predictive distributions and classical parametric posterior distributions are strongly inconsistent (incoherent); see Eaton and Sudderth (1995, 1999). The importance of the nonamenability of Gl_p in an applied context is discussed in Eaton, Muirhead and Pickering (2006). Further discussion of the nonamenability issue can be found in Kass and Wasserman (1996) and Eaton and Sudderth (2002, 2004).

The Stein (1956a) work provided troubling examples where traditional likelihood methods yield inferential procedures with some serious deficiencies. In essence, too much invariance yields incoherence. This circumstance has lead to a number of alternative proposals; none of which seem particularly appealing. The evidence would suggest that the foundational issues caused by nonamenability will not be resolved any time soon. This situation seems rather serious in multivariate analysis where Gl_p-invariance is ubiquitous.

4. The right Haar argument.   Fisher's fiducial inference stems in part from the observation that in certain situations, frequentist claims concerning confidence sets match probabilistic assertions based on a distribution on the parameter space. For example, suppose X has a univariate normal distribution with mean θ and variance one. Consider the set C that consists of the collection of pairs (x, θ) such that $|x - \theta| \leq d$ where d is a fixed positive constant. The set C has two sections

$$(6) \qquad\qquad C_x = \{\theta \mid (x, \theta) \in C\},$$

$$(7) \qquad\qquad C_\theta = \{x \mid (x, \theta) \in C\}.$$

Under the probability model for X, say $P(\cdot \mid \theta)$, the set C_θ has a fixed probability, say γ, that does not depend on the parameter θ. Next, let $Q(\cdot \mid x)$ be the normal distribution with mean x and variance one. Under this distribution for θ, the set C_x has the same probability γ for all x. In other words, the distribution Q for the parameter results in "probability matching" for the sections of the set C. The distribution on the parameter space arises via two different arguments, the first being what is called Fisherian pivoting. This pivoting uses the fact that the variable $(X - \theta)$ has a normal distribution with mean zero and variance one. Then with a wave of the hand, one asserts that the conditional distribution of θ given $X = x$ is just $Q(\cdot \mid x)$. The second argument assigns the parameter an improper prior distribution of Lebesgue measure and then formally applies Bayes theorem to obtain the posterior distribution $Q(\cdot \mid x)$ for the parameter. The fact that these two arguments agree in this and other relevant situations has long been of interest, especially after the Pitman work and the Stein (1965) and Hora and Buehler (1966) papers. An explanation for this agreement is given in Eaton and Sudderth (1999).

In the Hora and Buehler work, a statistical model that is invariant under a group G is assumed. Sufficiently limiting assumptions are made so that Fisherian pivoting can be used; see Eaton and Sudderth (1999) for the condition that yields Fisherian pivoting when G is transitive on the parameter space Θ. Hora and Buehler assume G equals Θ and focus on the fiducial distribution obtained from their structural assumptions; however, they also remark that the fiducial distribution is that obtained from a formal Bayes argument using right Haar measure as a prior distribution on $\Theta = G$. Hora and Buehler apply their results to derive best invariant estimators for certain parametric functions and to generalize some of the Pitman confidence set results. Stein (1965) considers an invariant model similar to that in Hora and Buehler, but the Stein considerations differ in two important aspects from both those in Hora and Buehler and earlier works. First, Stein does not make the standard structural assumptions needed to do Fisherian pivoting, and second, the focus is on the so-called right Haar prior distribution on Θ. The formal posterior distribution rather than the fiducial distribution becomes the center of attention. The importance of the Stein argument, although very sketchy, is that it suggests the validity of probability matching without Fisherian pivoting, and it hints at an alternative argument for finding best invariant decision rules. We now turn to a discussion of this alternative argument.

We begin the discussion of best invariant rules with a paraphrase of a result from Charles Stein circa 1965 [personal communication]: *Consider an invariant decision problem where the group G acts transitively on the parameter space. Using the right Haar prior distribution, calculate the formal posterior distribution on the parameter space. Then choose the action that minimizes the expected loss under the posterior distribution. This action, dependent upon the data, is a best invariant (minimum risk) decision rule.*

The exact provenance of this result is not known to us, as attempts to trace the origin have been largely unsuccessful. Zidek (1969), p. 292, calls the above "well known" but without a reference. Special cases are given in varying degrees of generality in a number of places. For example, see Peisakoff (1950), pp. 37–38, and Hora and Buehler (1966), Section 5. In Eaton (1970), the result is attributed to Stein and a proof is given. This proof uses a set of unappealing assumptions and utilizes the ideas and calculations in Stein (1965).

What is important about the Stein (1965) calculation is the emphasis on right Haar measure and the prior it induces on the parameter space. In cases where Fisherian pivoting is possible, this prior produces a posterior that is the fiducial distribution. The insight that the right Haar argument is more widely applicable than Fisherian pivoting has led to additional instances of probability matching examples (see Eaton and Sudderth (2002, 2004)). Further, the use of Haar measure has contributed significantly to a better understanding of the representation of best invariant rules described above (see Eaton (1989), Chapter 6, and Eaton and Sudderth (2001)).

5. Personal reflections. What a fortunate privilege it was for both of us to have studied under Charles Stein. The amazing depth of his ideas and imagination made a profound impression on our own thinking, one that would influence our research trajectories for years to come. But what was also so remarkable about this intellectual giant, were his wonderful human qualities, his kindness, humility and generosity to all. He would happily spend hours with us in his office, at lunch and on long walks, which he especially loved, discussing whatever we wanted, from statistics to the pressing political issues of the day. It is a true honor for us to contribute this memorial article to his legacy.

REFERENCES

BERGER, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR0804611 https://doi.org/10.1007/978-1-4757-4286-2

BONDAR, J. V. and MILNES, P. (1981). Amenability: A survey for statistical applications of Hunt–Stein and related conditions on groups. *Z. Wahrsch. Verw. Gebiete* **57** 103–128. MR0623457 https://doi.org/10.1007/BF00533716

BROWN, L. D. (1986). Commentary on paper [19]. In *Collected Papers of Jack Carl Kiefer. Supplementary Volume* (L. D. Brown, I. Olkin, J. Sacks and H. P. Wynn, eds.) 20–27. Springer-Verlag and IMS.

DEGROOT, M. H. (1986). A conversation with Charles Stein. *Statist. Sci.* **1** 454–462. MR0874478

EATON, M. L. (1970). Some problems in covariance estimation. Stanford University Technical Report No. 49. (This report is available at the website https://statistics.stanford.edu/resources/technical-reports).

EATON, M. L. (1989). *Group Invariance Applications in Statistics*. Institute of Mathematical Statistics, Beachwood, OH.

EATON, M. L. and FREEDMAN, D. A. (2004). Dutch book against some 'objective' priors. *Bernoulli* **10** 861–872. MR2093614 https://doi.org/10.3150/bj/1099579159

EATON, M. L., MUIRHEAD, R. J. and PICKERING, E. H. (2006). Assessing a vector of clinical observations. *J. Statist. Plann. Inference* **136** 3383–3414. MR2256273 https://doi.org/10.1016/j.jspi.2005.02.019

EATON, M. L. and SUDDERTH, W. D. (1995). The formal posterior of a standard flat prior in MANOVA is incoherent. *Journal of the Italian Statistical Society* **2** 251–270.

EATON, M. L. and SUDDERTH, W. D. (1999). Consistency and strong inconsistency of group-invariant predictive inferences. *Bernoulli* **5** 833–854. MR1715441 https://doi.org/10.2307/3318446

EATON, M. L. and SUDDERTH, W. D. (2001). Best invariant predictive distributions. In *Algebraic Methods in Statistics and Probability* (*Notre Dame*, *IN*, 2000) (M. A. G. Viana and D. S. P. Richards, eds.). *Contemp. Math.* **287** 49–62. Amer. Math. Soc., Providence, RI. MR1873666 https://doi.org/10.1090/conm/287/04776

EATON, M. L. and SUDDERTH, W. D. (2002). Group invariant inference and right Haar measure. *J. Statist. Plann. Inference* **103** 87–99. MR1896985 https://doi.org/10.1016/S0378-3758(01)00199-9

EATON, M. L. and SUDDERTH, W. D. (2004). Properties of right Haar predictive inference. *Sankhyā* **66** 487–512. MR2108204

GREENLEAF, F. P. (1969). *Invariant Means on Topological Groups and Their Applications. Van Nostrand Mathematical Studies*, *No.* 16. Van Nostrand Reinhold Co., New York-Toronto, Ont.-London. MR0251549

HORA, R. B. (1964). Fiducial probability theory for distributions with a group structure. Ph.D. thesis and Technical Report No. 33, Dept. Statistics, Univ. Minnesota.

HORA, R. B. and BUEHLER, R. J. (1966). Fiducial theory and invariant estimation. *Ann. Math. Stat.* **37** 643–656. MR0199938 https://doi.org/10.1214/aoms/1177699458

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

KASS, R. E. and WASSERMAN, L. (1996). The selection of prior distributions by formal rules. *J. Amer. Statist. Assoc.* **91** 1343–1370.

KIEFER, J. (1957). Invariance, minimax sequential estimation, and continuous time processes. *Ann. Math. Stat.* **28** 573–601. MR0092325 https://doi.org/10.1214/aoms/1177706874

KIEFER, J. (1966). Multivariate optimality results. In *Multivariate Analysis* (*Proc. Internat. Sympos.*, *Dayton*, *Ohio*, 1965) 255–274. Academic Press, New York. MR0226785

KUDO, H. (1955). On minimax invariant estimates of the transformation parameter. *Natur. Sci. Rep. Ochanomizu Univ.* **6** 31–73. MR0077842

LEHMANN, E. L. (1950). Some principles of the theory of testing hypotheses. *Ann. Math. Stat.* **21** 1–26. MR0034001 https://doi.org/10.1214/aoms/1177729884

LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*, 1st ed. Wiley, New York. MR0107933

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

OLKIN, I. and SELLIAH, J. B. (1977). Estimating covariances in a multivariate normal distribution. In *Statistical Decision Theory and Related Topics*, *II* (*Proc*. *Sympos*., *Purdue Univ*., *Lafayette*, *Ind*., 1976) (S. S. Gupta and D. S. Moore, eds.) 313–326. Academic Press, New York. MR0436407

PEISAKOFF, M. P. (1950). Transformation parameters. Ph.D. dissertation from Princeton Univ., unpublished.

PITMAN, E. J. G. (1939a). The estimation of the location and scale parameters of a continuous population of any given form. *Biometrika* **30** 391–421.

PITMAN, E. J. G. (1939b). Tests of hypotheses concerning location and scale parameters. *Biometrika* **31** 200–215. MR0000382 https://doi.org/10.1093/biomet/31.1-2.200

ROBERT, C. P. (2007). *The Bayesian Choice*: *From Decision-Theoretic Foundations to Computational Implementation*, 2nd ed. *Springer Texts in Statistics*. Springer, New York.

SELLIAH, J. B. (1964). Estimation and testing problems in a Wishart distribution. Ph.D. thesis and Technical Report No. 10, Dept. Statistics, Stanford Univ.

STEIN, C. (1956a). Some problems in multivariate analysis, Part I. Stanford University Technical Report No. 6. (This report is available at the website https://statistics.stanford.edu/resources/technical-reports).

STEIN, C. (1956b). 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). An example of wide discrepancy between fiducial and confidence intervals. *Ann*. *Math*. *Stat*. **30** 877–880. MR0125680 https://doi.org/10.1214/aoms/1177706072

STEIN, C. (1965). Approximation of improper prior measures by prior probability measures. In *Proc*. *Internat*. *Res*. *Sem*., *Statist*. *Lab*., *Univ*. *California*, *Berkeley*, *Calif*., 1963 (J. Neyman and L. LeCam, eds.) 217–240. Springer, New York. MR0199937

VON NEUMAN, J. (1929). Zur allgemeinen theorie des masses. *Fund*. *Math*. **13** 73–116.

WALD, A. (1939). Contributions to the theory of statistical estimation and testing hypotheses. *Ann*. *Math*. *Stat*. **10** 299–326. MR0000932 https://doi.org/10.1214/aoms/1177732144

WALD, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Trans*. *Amer*. *Math*. *Soc*. **54** 426–482. MR0012401 https://doi.org/10.2307/1990256

ZIDEK, J. V. (1969). A representation of Bayes invariant procedures in terms of Haar measure. *Ann*. *Inst*. *Statist*. *Math*. **21** 291–308. MR0246431 https://doi.org/10.1007/BF02532257