

REFERENCES

1. H. B. Cohen, *Injective envelopes of Banach spaces*, Bull. Amer. Math. Soc. **70** (1964), 723–726.
2. M. Coste, *Une approche logique des théories définissables par limites projectives finies*, Séminaire Benabou, Univ. Paris-Nord, 1976.
3. C. Ehresmann, *Gattungen von lokalen Strukturen*, Jahresber. Deutsch. Math.-Verein. **60** (1957), 59–77.
4. M. P. Fourman and J. M. E. Hyland, *Sheaf models for analysis*, Lecture Notes in Math., vol. 753, Springer-Verlag, 1979, pp. 280–301.
5. G. Gierz, et al, *A compendium of continuous lattices*, Springer-Verlag, 1980.
6. A. M. Gleason, *Projective topological spaces*, Illinois J. Math. **2** (1958), 482–489.
7. D. B. Goodner, *Projections in normed linear spaces*, Trans. Amer. Math. Soc. **69** (1950), 89–108.
8. A. Grothendieck and J. Dieudonné, *Éléments de géométrie algébrique*. I, Inst. Hautes. Études Sci. Publ. Math. No. 4 (1960).
9. M. Hasumi, *The extension property of complex Banach spaces*, Tôhoku Math. J. **10** (1958), 135–142.
10. M. Hochster, *Prime ideal structure in commutative rings*, Trans. Amer. Math. Soc. **142** (1969), 43–60.
11. J. Isbell, *Three remarks on injective envelopes of Banach spaces*, J. Math. Anal. Appl. **27** (1969), 516–518.
12. ———, *General functorial semantics*. I, Amer. J. Math. **94** (1972), 535–596.
13. ———, *Some concrete dualities*, Notices Amer. Math. Soc. **21** (1974), A567–A568.
14. R. Kaufman, *A type of extension of Banach spaces*, Acta Sci. Math. (Szeged) **27** (1966), 163–166.
15. J. L. Kelley, *Banach space with the extension property*, Trans. Amer. Math. Soc. **72** (1952), 323–326.
16. L. Nachbin, *A theorem of the Hahn-Banach type for linear transformations*, Trans. Amer. Math. Soc. **68** (1950), 28–46.
17. R. S. Phillips, *On linear transformations*, Trans. Amer. Math. Soc. **48** (1940), 516–541.
18. J. Rainwater, *A note on projective resolutions*, Proc. Amer. Math. Soc. **10** (1959), 734–735.
19. Z. Semadeni, *Projectivity, injectivity, and duality*, Rozprawy Mat. **35** (1963).
20. R. Sikorski, *A theorem on extensions of homomorphisms*, Ann. Soc. Polon. Math. **21** (1948), 332–335.

JOHN ISBELL

BULLETIN (New Series) OF THE
 AMERICAN MATHEMATICAL SOCIETY
 Volume 11, Number 2, October 1984
 © 1984 American Mathematical Society
 0273-0979/84 \$1.00 + \$.25 per page

Statistical estimation, asymptotic theory, by I. A. Ibragimov and R. Z. Has'minskii, Applications of Mathematics, vol. 16, Springer-Verlag, New York, 1981, vii + 403 pp., \$42.00. ISBN 0-3879-0523-5

Contributions to a general asymptotic statistical theory, by J. Pfanzagl (with the assistance of W. Wefelmeyer), Lecture Notes in Statistics, vol. 13, Springer-Verlag, New York, 1982, vii + 315 pp., \$16.80. ISBN 0-3879-0776-9

Asymptotic statistical theory is a body of limit or, better yet, approximation theorems used by statisticians to elude the intractability of all but the very simplest practical statistical problems and to obtain usable results. As such it is not a subject with a well-defined scope or natural boundaries. An early

contributor to the subject was Laplace. He may have had predecessors, but I did not know them personally. At any rate, in his papers published in 1809–1810, Laplace did present a number of interesting results. He gave a fairly general formulation of what is now called the Central Limit Theorem, with a proof that is certainly valid for sums of independent bounded lattice variables and can be extended further. Laplace also proved that for smoothly parametrized families of distributions, as the number of observations tends to infinity, posterior distributions tend to Gaussian distributions centered on what Fisher was to call the maximum likelihood estimates. Laplace used that to prove asymptotic optimality properties of Bayes (or m.l.e.) estimates and to give a justification for Legendre's method of least squares.

The statistical aspects of the subject seem to have remained fairly dormant until they were revived by R. A. Fisher in 1922 and 1925. He introduced names, such as "maximum likelihood", "efficiency", "consistency" and carried out a bit further some of Laplace's arguments. Fisher spurned the use of Bayes' theorem and did not recover those of Laplace's results that are connected with it.

There is now a very large body of results about problems that can be described roughly as follows.

Take a σ -field \mathcal{A} of subsets of a set \mathcal{X} . Let \mathcal{F} be a family of probability measures on \mathcal{A} and let t be a "functional", that is, a function from \mathcal{F} to the line, or a Euclidean space, or perhaps some other space. Let X_1, X_2, \dots, X_n be independent identically distributed observations whose individual distribution is some measure $P \in \mathcal{F}$. A statistic T_n is a function from \mathcal{X}^n to the range of t . The asymptotic statistician will study the limiting behavior of the distribution of T_n as n tends to infinity. If T_n is intended as an estimate of t , one indicates how good or bad T_n is by the use of loss functions and their expectations, just as was done by Laplace. Pfanzagl and Wefelmeyer (henceforth abbreviated [P.W.]) prefer to use measures of concentration of the distribution of T_n around t (see their comments, p. 151 s.q.q.). If T_n is intended for testing purposes, one attempts to obtain approximate values for the power functions of the Neyman-Pearson theory.

A large effort has been spent on functions T_n obtained through the maximum likelihood method, perhaps because Fisher claimed that these were always the optimal ones. Other favorites are the von Mises differentiable functions, Hoeffding's U -statistics, the statistics provided by minimum distance methods, and the statistics "based on ranks". The last have boomed particularly in the fifties and sixties after it was shown by E. Lehmann and others that they do have rather remarkable properties (see Hájek-Šidák (1967) and E. Lehmann (1975)).

Here we have mentioned only i.i.d. observations, for simplicity, and because this is the domain that has yielded the largest volume of papers. However there is a massive amount of material available on other cases. The Ibragimov-Has'minskii text (henceforth abbreviated [I.H.]) contains a bit of it for independent nonidentically distributed observations and for special stochastic processes. Pfanzagl limits himself essentially to the i.i.d. case.

Another important part of the enterprise is connected with the name “robustness”. There one investigates what happens to the statistics T_n if the true distribution of the X_j is not one of the postulated family \mathcal{F} .

According to [P.W.]’s Introduction, these endeavors are based on thin logic, and they lack coherence. This is certainly true and probably all for the best: no logically organized theory, conceived in the present, can be expected to cope with major future developments that are not yet perceived.

It is, however, possible to give a good portion of the field a solid foundation and some organization if one sticks by the precepts of the Laplace-Neyman-Pearson-Wald statistical decision theory, or close to them. The main idea is simple. It will be described briefly since it is not explicit in [I.H.] or [P.W.].

Call any family $\mathcal{E} = \{P_\theta: \theta \in \Theta\}$ of probability measures P_θ in a σ -field \mathcal{A} an “experiment indexed by a set Θ ”.

Let $\mathcal{E} = \{P_\theta: \theta \in \Theta\}$ and $\mathcal{E}' = \{P'_\theta: \theta \in \Theta\}$ be two experiments with the same index set Θ but with possibly different underlying spaces $(\mathcal{X}, \mathcal{A})$ and $(\mathcal{X}', \mathcal{A}')$. It turns out that we can define a number $\Delta = \Delta(\mathcal{E}, \mathcal{E}')$ that measures the “distance” between \mathcal{E} and \mathcal{E}' for several different purposes. One of them is that for any loss function W such that $0 \leq W \leq 1$, any risk function achievable on one of the experiments can be matched within Δ by a risk function achievable on the other experiment. Another interpretation of Δ is that a randomization operation conducted after carrying out \mathcal{E} can reproduce the measures of \mathcal{E}' within 2Δ for the total variation norm, and similarly, \mathcal{E} and \mathcal{E}' being interchanged. The principle of the operation is then to replace an intractable \mathcal{E} by a simpler \mathcal{E}' within short distance of \mathcal{E} , treat the problem there, and return to \mathcal{E} by the randomization operation.

An idea very similar to this occurs in Wald’s fundamental paper of 1943. There Wald looks at experiments $\mathcal{E}^n = \{P_\theta^n: \theta \in \Theta\}$ formed by the distribution P_θ^n of n identically distributed observations. He shows that under sufficient regularity assumptions one can replace \mathcal{E}^n by an experiment $\mathcal{F}_n = \{Q_{\theta,n}: \theta \in \Theta\}$, where $Q_{\theta,n}$ is the Gaussian approximation to the distribution of the maximum likelihood estimate.

It was noted some years later that Wald’s argument depends mainly on two features of his special situation: (a) the existence of sufficiently accurate estimates, and (b) the possibility of local approximation of the experiment by a Gaussian shift one. This led to the introduction of the locally asymptotically normal families (LAN), to the asymptotic minimax theorem sometimes called by the names Hájek and Le Cam, to Hájek’s celebrated convolution theorem and a number of other “general” statements. For extensions to the “locally asymptotically mixed normal” (LAMN) families so prevalent in the study of stochastic processes, see the book by Basawa and Prakasa Rao (1980), the work of Jeganathan (1980), and the notes by Basawa and Scott (1983).

The Hájek-Le Cam asymptotic minimax theorem says that if a risk function is not achievable on an experiment \mathcal{E} , it is eventually not achievable either along a sequence $\{\mathcal{E}_n\}$ of experiments that tend to \mathcal{E} , even in a weak way. This means, for instance, that if the limit is Gaussian one can read off lower bounds for limits of risks along the sequence from those already known in the Gaussian case. The Hájek convolution theorem involves statistics T_n with values in a group G and with limiting distributions F_θ permuted among each

other by the action of G . In a variety of cases there are “distinguished” sequences $\{T_n\}$ with the said property and the further property that if $\{T'_n\}$ is any other sequence with limiting distributions F'_θ similarly permuted by G then F'_θ is the convolution of F_θ by some probability measure μ that depends on $\{T'_n\}$ but not on θ . This was proved by Hájek for LAN families with Euclidean parameter sets taken for the group G . Later Moussatat (1976) and Millar (1982) extended it to LAN families parametrized by the additive group G of a Hilbert space. This reviewer considered some other cases.

A combination of these relatively simple results with the ideas of C. Stein (1956) about the influence of nuisance parameters gives a connecting thread through a good part of asymptotic statistical theory. It does not cope with everything. Left out is the excellent work of Bahadur. This uses “large deviations” with probabilities that are (asymptotically) too small to be of importance when one tries to approximate experiments. The use of “large deviations” is often connected with a squabble about whether one should consider “fixed” instead of “near-by” alternatives. This is due to standard presentations where everything is “fixed” except the number n of observations which tends to infinity. The silliness of the question is readily apparent if one thinks of approximations instead of limits, or, as this reviewer likes to do, make everything including the parameter space depend on n .

Besides results that involve large deviations, there is a very substantial body of theory that uses “asymptotic expansions” in the usual mathematical sense instead of the mere first order approximations that were sufficient for Laplace and some of his successors. (See for instance Akahira and Takeuchi (1981), Pfanzagl (1980).)

The two books under review do not deal with these higher order expansions. Pfanzagl intends to devote a second volume to them.

The purpose of the [I.H.] book is to present certain aspects of asymptotic theory of estimation with emphasis on the theoretical mathematical aspects. The instances selected for specific study involve mostly parametrized families in the independent identically distributed case and signal + noise examples from communication theory. There are, however, a substantial section on independent nonidentically distributed observations and a chapter on non-parametric problems. The first chapter contains a variety of results about the estimation problem: construction of consistent estimates, Cramér-Rao inequalities, Pitman estimates, Bayes estimates, maximum likelihood, method of moments, etc. This first chapter by itself contains more than the usual course or textbook contents on the subject. Then [I.H.] proceeds to study the LAN assumptions and their consequences. Two further chapters are devoted to “irregular” densities where the limit experiments are Poisson instead of Gaussian ones. Chapter VII deals with estimation for stochastic processes with Gaussian noise.

Apparently [P.W.]’s purpose was to present a general methodological framework illustrated by some results that involve exclusively Gaussian approximations and mostly independent identically distributed observations. The basic aim is to deal with estimation of functionals defined on general nonparametric families. (This reviewer found Pfanzagl’s own description of purpose, pp. 8 and 9, rather mysterious. Thus the above may be wrong.) The statistical part of the

book, beginning with Chapter 8, p. 115, actually contains very few “general” results. It does contain lower bounds on the risk of estimates, a form of Hájek’s convolution theorem, and bounds on risks related to Stein’s ideas. The non-parametric context of [P.W.] makes it more “general” than [I.H.]. For parametric families the LAN context of [I.H.] goes further than the i.i.d. case of [P.W.].

A particular feature of [P.W.] is the great emphasis on “tangent spaces” in Chapters 1 to 7. Pfanzagl and Wefelmeyer work with a particular probability measure P_0 and the Hilbert space generated by a family \mathcal{F} of probability densities that are P_0 -square integrable. They do have a definition of “weak differentiability” that avoids the restriction to square integrable densities (p. 23), but that is used only sporadically in the sequel. It was shown by this reviewer, with input from Pfanzagl, that the weak differentiability is equivalent to the Hellinger differentiability used by Beran, Bickel, Hájek, Koshevnik, Levit, Millar and many others, including this reviewer.

The relevance of tangent spaces obtained from Hellinger distances can be seen more clearly as follows. Consider an experiment $\mathcal{G} = \{G_\theta; \theta \in \Theta\}$ and a particular $t \in \theta$. Call \mathcal{G} a Gaussian shift experiment if the G_θ are mutually absolutely continuous and if the stochastic process

$$\theta \rightarrow X'_\theta = \log \frac{dG_\theta}{dG_t},$$

with the distributions induced by G_t , is a Gaussian process. Such an experiment generates a Hilbert space \mathcal{H} , the closed linear span of the process

$$\theta \rightsquigarrow X_\theta = X'_\theta - E_t X'_\theta.$$

If, as we shall assume for simplicity, $\theta' \neq \theta''$ implies $G_{\theta'} \neq G_{\theta''}$, the map $\theta \rightsquigarrow X_\theta$ imbeds Θ in \mathcal{H} . Then the statistical properties of the experiment \mathcal{G} are entirely determined by the metric structure of Θ as a subset of \mathcal{H} . Conversely, any subset of a Hilbert space has a standard Gaussian shift experiment attached to it. Now note that finite positive measures on a σ -field \mathcal{A} can be imbedded in a Hilbert space \mathcal{H}_n using the square norm $4nf(\sqrt{dp} - \sqrt{dq})^2$. This way a set $\{p_\theta; \theta \in \Theta_n\}$ of probability measures on \mathcal{A} yields a Gaussian shift experiment \mathcal{G}_n . It also yields a product experiment $\mathcal{E}_n = \{p_\theta^n; \theta \in \Theta_n\}$, where p_θ^n is the distribution of n i.i.d. observations from p_θ . In the local case, that is, for measures p_θ^n that do not separate entirely, it turns out that when approximation by a Gaussian shift experiment is possible, then, under mild restrictions, it can be achieved through the particular \mathcal{G}_n described above.

This explains in part the relevance of the local Hilbert structure of Θ_n for the Hellinger distance $h(P, Q)$ defined by $h^2(P, Q) = \frac{1}{2}f(\sqrt{dP} - \sqrt{dQ})^2$. Another main reason for the use of h is that it is easily computable on direct products from its values on components:

$$1 - h^2 \left[\left(\prod_j P_j \right), \left(\prod_j Q_j \right) \right] = \prod_j [1 - h^2(P_j, Q_j)].$$

Instead of the Hellinger distance, [P.W.] uses, locally, square distances of the chi-square type:

$$\Delta^2(P, Q; R) = \int \frac{(dP - dQ)^2}{dR}.$$

(The corresponding inner products were also used by Hellinger.) This can be done under some restrictions. In particular, note that the square distance

$$K^2(P, Q) = \frac{1}{2} \int \frac{(dP - dQ)^2}{d(P + Q)}$$

satisfies $h^2 \leq k^2 \leq 2h^2$, and, under some regularity restrictions, k^2 behaves like $2h^2$ for h small. However, such distances do not propagate well when one takes direct products. [P.W.] used such distances to introduce the material in a simpler manner. Unfortunately, this means quite a bit of additional complexity in Chapters 6 and 7, for instance.

It should be clear from this that the tangent spaces of [P.W.] are relevant in that they allow at least some description of a very relevant local Hilbertian structure. Note, however, that this is a special feature of the i.i.d. case. It is not adequate in a “general asymptotic statistical theory” that should be able to cope with regression problems, time series or the signal + noise schemes of [I.H.].

The LAN assumptions of Le Cam (1960) involve two main features. One of them is local approximability by Gaussian shift experiments \mathcal{G}_n . The other assumes that the parameter spaces are subsets of a Euclidean space \mathbf{R}^k and that the Hilbert spaces of the \mathcal{G}_n can be represented linearly in \mathbf{R}^k . In such a system the tangent spaces introduced in statistics by H. Chernoff (1954) are exactly what is needed to allow local replacement of a set by a tangent space. In this respect note that [P.W.] introduces two concepts: one called “tangent cone” (p. 23), and a modification (p. 24). This latter is the “contingent” used by Bouligand (1932). See also Saks (1937).

There are cases where the “tangent cone” is reduced to the origin, but where the contingent is the entire space. There are also cases where neither of these objects satisfy Chernoff’s requirements. Thus they should be used with caution, even in the i.i.d. case and even if they are taken for Hellinger distances.

There are other surprises in [P.W.]. For instance Chapters 10 and 11 have titles that begin “Existence of Asymptotically Efficient Estimates”, but they do not actually provide conditions under which the estimates exist or proof of existence except in particular examples. [P.W.] derides the use of minimum distance methods, but does not tell us how to obtain the estimators that he improves in Chapter 11.

In spite of all of this, [P.W.] is well worth reading, would it be only for its wealth of good examples. At places it is also enlivened by diatribes about some common practices. Some of them are in Chapter 0 but see also p. 120 s.q.q. and 151 s.q.q. One does not have to agree with the sentiments expressed there to enjoy reading them.

The [I.H.] book is also enjoyable reading except perhaps for one feature: there are many proofs of convergence of distributions of stochastic processes. Such proofs tend to be complex, because such is the nature of the beast. The results are not really needed for the main statistical implications, although they are needed for the treatment of particular estimates, such as maximum likelihood.

There are some minor matters that left this reader feeling like finding blemishes on a well-polished apple. I did not find any mention of the fact that Bayes estimates can be disastrously inconsistent (see Freedman (1963)) or any warning of that nature. Another item is the choice of formulation for the LAN assumptions (p. 120). They involve a particular point $t \in \mathbf{R}^k$ and nearby points of the form $t + \varphi(\varepsilon)u$, where φ is a matrix and u is an element of \mathbf{R}^k . In the [I.H.] formulation, the convergence to zero of the remainder term $\psi_\varepsilon(u, t)$ is assumed to take place pointwise in u . This is indeed the version used by Hájek in (1970) and (1972). The original version of Le Cam (1960) required that $\psi_\varepsilon(u_\varepsilon, t)$ tend to zero as long as u_ε remains bounded. Hájek's version is sufficient to obtain convolution theorems and a minimax lower bound. It is not sufficient to insure attainability of the lower bound.

It is true that [I.H.] gives another definition (p. 123). This one requires uniformity of the convergence in u , but also in t , an uncomfortable mixture.

Another item is the form of the statement of the asymptotic minimax theorems. [I.H.] take the supremum of the risk over neighborhoods $|\theta - t| \leq \delta$. This is indeed the way the theorem is stated in Hájek (1972), p. 186. However, it is proved there for shrinking neighborhoods, as stated in that same paper, p. 189, Remark 1. For the special case considered at that time, Le Cam (1953) also used shrinking neighborhoods. The use of the neighborhoods $|\theta - t| \leq \delta$ was roundly criticized in Fabian and Hannan (1982). It is perhaps not a point of major importance except that most applications to nonparametric situations seem to need the shrinking neighborhood version. See for instance [I.H.], p. 238, where [I.H.] reproduces a result of Roger Farrell (mistransliterated throughout as Farrel).

In connection with this result, it would have been nice to point out that the technique of proof does not apply to situations where one estimates the entire density with a loss function such as the Hellinger distance (see Remark 5.5, p. 140). The [I.H.] proof of Farrell's result relies on the fact that one can modify a density at a point very materially and still keep very close for the Hellinger distance. For other problems one needs to make a stronger use of the dimensionality of the parameter space. There are results of that nature. They rely either on Fano's lemma ([I.H.], p. 323) or on P. Assouad's lemmas (1983); see also Birgé (1983).

Chapters V and VI could have been made more transparent by a systematic use of the fact that it is enough to prove the desired convergences for experiments where the number of observations is a random number N , independent of everything else, with a Poisson distribution such that $EN = n$.

The discussion of the properties of estimates in VI, §6 could have included remarks to the effect that the estimates do often depend materially on the choice of loss functions, contrary to what happens in the LAN situation, for instance.

Since this is a translation, one would have expected a scarcity of typographical errors. Unfortunately the translation has provided a few that were not in the original. This reviewer did not attempt to search for them, but, for instance, the Legendre duplication formula for gamma functions on p. 306 is missing two of the gammas present in the original. There are also some

peculiar translations. For instance, p. 128: . . . “Indeed Lindeberg’s theorem as well as the theorem concerning the relative stability of some random variables remain valid when applied to the sequences of the series”. Presumably this means that Lindeberg’s theorem (and something else that is difficult to identify) remains valid for triangular arrays. Another example occurs in Fano’s lemma (p. 323): “Let $\Psi(X)$ be a decision rule (an estimator) which corresponds to each ‘observation’ X of the values $\theta_1, \theta_2, \dots, \theta_n$.” I think they mean “assigns to each X one of the θ_i ”. There are many more.

The [I.H.] book would serve well as a graduate level introduction to the subject. It is also a good reference for the territory it covers. For serious readers it should be supplemented since it gives only a view of a small part of a very large domain.

Some related works are the Basawa-Prakasa Rao book already mentioned, R. Serfling’s book on *Approximation theorems*, U. Grenander’s *Abstract inference* and the *Martingale limit theory* of Hall and Heyde.

The [P.W.] book appears in a series that advertises itself by saying that “The timeliness of a manuscript is more important than its form which may be unfinished or tentative”. A caution to this effect is also given on p. 20 of [P.W.]. If “form” includes typography, one can indeed hope that a more polished version will become available. The typography even bothered some of my younger colleagues who do not have the excuse of weak eyesight.

There is room for improvement and much “unfinished business”, especially in Chapters 10 and 11—this is due to the present state of the art. One of the main impacts of [P.W.] is that it has already prodded several authors to improve the state of the art. It will certainly continue to inspire many more.

REFERENCES

- M. Akahira and K. Takeuchi (1981), *Asymptotic efficiency of statistical estimators: Concepts and higher order asymptotic efficiency*, Lecture Notes in Statistics, vol. 7, Springer-Verlag, Berlin and New York.
- P. Assouad (1983), *Deux remarques sur l’estimation*, C. R. Acad. Sci. Paris, **296**, 1021–1024.
- I. Basawa and B. L. S. Prakasa Rao (1980), *Statistical inference for stochastic processes*, Academic Press, New York.
- I. Basawa and D. J. Scott (1983), *Asymptotic optimal inference for non-ergodic models*, Lecture Notes in Statistics, vol. 17, Springer-Verlag, Berlin and New York.
- R. Beran (1974), *Asymptotically efficient adaptive rank estimates in location models*, Ann. Statist. **2**, 63–74.
- P. Bickel (1982), *On adaptive estimation*, Ann. Statist. **10**, 647–671.
- L. Birgé (1983), *On estimating a density using Hellinger distance and some other strange facts*, MSRI Report 045–83.
- G. Bouligand (1932), *Introduction à la géométrie infinitésimale directe*, Gauthier-Villars, Paris.
- H. Chernoff (1954), *On the distribution of the likelihood ratio*, Ann. Math. Statist. **25**, 573–578.
- V. Fabian and J. Hannan (1982), *On estimation and adaptive estimation for locally asymptotically normal families*, Z. Wahrsch. Verw. Gebiete **59**, 459–478.
- R. A. Fisher (1922), *On the mathematical foundations of theoretical statistics*, Philos. Trans. Roy. Soc. London Ser. A. **222**, 309–368.
- _____ (1925), *Theory of statistical estimation*, Proc. Cambridge Philos. Soc. **22**, 700–725.
- D. A. Freedman (1963), *On the asymptotic behavior of Bayes estimates in the discrete case*, Ann. Math. Statist. **34**, 1386–1403.
- U. Grenander (1981), *Abstract inference*, Wiley, New York.

J. Hájek (1972), *Local asymptotic minimax and admissibility in estimation*, Proc. Sixth Berkeley Sympos. Math. Statist. and Probab. Vol. 1, pp. 175–194.

J. Hájek and Z. Šidák (1967), *Theory of rank tests*, Academia, Praha.

P. Hall and C. C. Heyde (1980), *Martingale limit theory and its application*, Academic Press, New York.

P. Jeganathan (1980), *Asymptotic theory of estimation when the limit of the loglikelihood ratios is mixed normal*, Thesis, Indian Statist. Inst. Calcutta.

Yu. A. Koshvnik and B. Ya. Levit (1976), *On a non-parametric analogue of the information matrix*, Theory Probab. Appl. **21**, 738–753.

P. Laplace (1809–1810), *Approximation des formules qui sont fonctions de très grands nombres et leur application aux probabilités*, Mémoires de l'Académie des Sciences, vol. 10.

L. Le Cam (1960), *Locally asymptotically normal families of distributions*, Univ. California Publ. Statist., vol. 3, pp. 37–98.

E. L. Lehmann (1975), *Nonparametrics. Statistical methods based on ranks*, Holden-Day, San Francisco, Calif.

P. W. Millar (1979), *Asymptotic minimax theorems for the sample distribution function*, Z. Wahrsch. Verw. Gebiete **48**, 233–252.

_____ (1983), *Nonparametric applications of an infinite dimensional convolution theorem*, Preprint, Dept. Statist., Berkeley.

W. Moussatat (1976), *On the asymptotic theory of statistical experiments and some of its applications*, Thesis, Univ. of California, Berkeley.

J. Pfanzagl (1980), *Asymptotic expansions in parametric statistical theory*, Developments in Statistics (P. R. Krishnaiah, ed.), Academic Press, New York, pp. 1–97.

R. Serfling (1982), *Approximation theorems of mathematical statistics*, Wiley, New York,

S. Saks (1937), *Theory of the integral*, 2nd ed., Hafner, New York.

C. Stein (1956), *Efficient nonparametric testing and estimation*, Proc. Third Berkeley Sympos. Math. Statist. and Probab., Vol. 1, pp. 187–195.

A. Wald (1943), *Tests of statistical hypotheses concerning several parameters when the number of observations is large*, Trans. Amer. Math. Soc. **54**, 426–482.

LUCIEN LE CAM

BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 11, Number 2, October 1984
© 1984 American Mathematical Society
0273-0979/84 \$1.00 + \$.25 per page

Manifolds all of whose geodesics are closed, by Arthur L. Besse, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 93, Springer-Verlag, Berlin-Heidelberg-New York, ix + 262 pp., 1978, \$40.00. ISBN 0-3870-8158-5

In differential geometry, as in many branches of mathematics, the practitioners can be classified roughly into two groups, the structuralists and the problem solvers. Flourishing in the time of Hilbert and reaching a peak sometime after the appearance of the Bourbaki series, the structuralists gained the upper hand. However, the titles of books recently published, such as *Comparison theorems in Riemannian geometry* by J. Cheeger and D. Ebin, North-Holland, 1975, and the book at hand seem to indicate that the rococo in mathematics, especially differential geometry, has come back in force.

If we probe a little deeper, we find that the relation between the two schools is more cooperative than competitive. Some of the problems discussed in this