formly in λ. However, this approach may not succeed in controlling the asymptotic *level* of the confidence set. The problem is that the relevant asymptotic expansions may not converge uniformly over all λ, especially when λ is infinite dimensional.

An interesting strategy, proposed by Loh (1985) in a testing context, is to pick critical values so as to control the apparent level of the confidence set for θ over a confidence set for λ of level $1 - \varepsilon_n$, where ε_n is small. When feasible, this construction ensures that the level of the confidence set for θ is at least $1 - \alpha - \varepsilon_n$. One difficulty is finding a good confidence set for λ. If the latter is too large, then the induced confidence set for θ is likely to be inefficient. Perhaps the notion of controlling level of a confidence set for θ is too strong. On the other hand, controlling asymptotic coverage probability only pointwise in λ is clearly too weak.

REFERENCES

BERAN, R. J. (1987). Prepivoting to reduce level error of confidence sets. *Biometrika* **74** 457–468.

BERAN, R. J. (1988). Prepivoting test statistics: A bootstrap view of asymptotic refinements. *J. Amer. Statist. Assoc.* To appear.

BICKEL, P. J. (1987). Discussion of Efron. *J. Amer. Statist. Assoc.* **82** 191.

DUCHARME, G. R., JHUN, M., ROMANO, J. P. and TRUONG, K. N. (1985). Bootstrap confidence cones for directional data. *Biometrika* **72** 637–645.

LOH, W.-Y. (1985). A new method for testing separate families of hypotheses. *J. Amer. Statist. Assoc.* **80** 362–368.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720

PETER J. BICKEL

University of California, Berkeley

I found Professor Hall's unified treatment of bootstrap bounds and confidence intervals very valuable. I was particularly interested by his exploration of the relation between the accelerated bias correction bootstrap bounds and the Studentized bounds, a relationship which I also studied, but only in the parametric framework, in my discussion of Efron (1987). In my discussion I want to:

1. Argue at least heuristically that, in the nonparametric context, the second order equivalence of $\hat{\theta}_{\mathrm{ABC}}$ and $\hat{\theta}_{\mathrm{STUD}}$ holds quite generally for $\theta(F)$ a sufficiently smooth von Mises functional, provided that we Studentize properly. For example, it holds if the estimate $\hat{\theta} = \theta(\hat{F})$, where $\hat{F}$ is the empirical d.f., is an M estimate corresponding to a nice ψ function; see Huber [(1981), Chapter 2] for examples.
2. Suggest that quite generally in a parametric, nonparametric or semiparametric context, $\hat{\theta}_{\mathrm{ABC}}$ and $\hat{\theta}_{\mathrm{STUD}}$ are second order equivalent provided again that $\hat{\theta}$ is efficient and we Studentize properly, that is, by an efficient estimate of the asymptotic standard deviation of $\hat{\theta}$.

Since, modulo regularity conditions, checking equivalence of the theoretical points $\hat{\theta}_{\text{abc}}$ and $\hat{\theta}_{\text{Stud}}$ and of the corresponding bootstrap points $\hat{\theta}_{\text{AB}}$ and $\hat{\theta}_{\text{STUD}}$ is the same thing, we restrict ourselves to the former. Suppose H and K are as in Hall the d.f.'s of $T \equiv \sqrt{n}(\hat{\theta} - \theta(F))/(\sigma(F))$ and $\tilde{T} \equiv \sqrt{n}(\hat{\theta} - \theta(F))/\hat{\sigma}$ and admit Edgeworth expansions,

$$H(x) = \Phi(x) - \phi(x)\left(K_{1n} + \frac{K_{3n}}{6}(x^2 - 1)\right) + O(n^{-1}),$$

$$K(x) = \Phi(x) - \phi(x)\left(\tilde{K}_{1n} + \frac{\tilde{K}_{3n}}{6}(x^2 - 1)\right) + O(n^{-1}),$$

where $K_{in}, \tilde{K}_{in} = O(n^{-1/2})$, $i = 1, 3$, are asymptotic cumulants of $T, \tilde{T}$; see Bickel (1974) for a precise formulation. In this notation, Hall shows that $\hat{\theta}_{\text{abc}} = \hat{\theta}_{\text{Stud}} + O(n^{-3/2})$ provided that

$$(1) \qquad m = \Phi^{-1}H(0) + O(n^{-1}) = K_{1n} - \frac{K_{3n}}{6} = \tilde{K}_{1n} - \frac{\tilde{K}_{3n}}{6} + O(n^{-1})$$

and

$$(2) \qquad a = -\frac{1}{6}\left(K_{3n} + \tilde{K}_{3n}\right) + O(n^{-1}).$$

Note that without any further calculation, the corresponding quantities for $\hat{\theta}_{\text{ABC}}$, not only $\hat{m}$ but also $\hat{a}$ can be estimated (but not in an invariant way) directly. In particular, we can take $\hat{a} = -\frac{1}{6}(\hat{K}_{3n} + \hat{\tilde{K}}_{3n})$, where the $K_{3n}, \tilde{K}_{3n}$ are the bootstrap cumulants of appropriately truncated versions of T and $\tilde{T}$. For the fully nonparametric model and special $\theta(F)$ and exponential families Hall, following Efron (1987), derives a further approximation for a, if $\hat{\theta}$ is an efficient estimate of $\theta(F)$ and $\hat{\sigma}$ is an efficient estimate of $\sigma(F)$, the asymptotic s.d. of $\sqrt{n}\,\hat{\theta}$.

To extend Hall's approach in the fully nonparametric model, suppose $\hat{\theta}$ admits a von Mises (stochastic) expansion

$$\hat{\theta} = \theta(F) + \int \psi(x, F)\, d\hat{F}(x) + \tfrac{1}{2}\int \gamma(x, y, F)\, d\hat{F}(x)\, d\hat{F}(y) + o_P(n^{-1}),$$

where

$$\int \Phi(x, F)\, dF(x) = 0,$$

γ is symmetric and

$$\int \gamma(x, y, F)\, dF(y) = 0 \quad \text{for all } x.$$

Compare Withers (1983). The efficient estimate of $\sigma^2(F) = \int \psi^2(x, F)\, dF(x)$ is $\sigma^2(\hat{F}) = \int \psi^2(x, \hat{F})\, d\hat{F}(x)$. Under additional regularity conditions we can check $\sigma^2(\hat{F}) = \sigma^2(F) + \int v(x, F)\, d\hat{F} + O_P(n^{-1})$, where

$$v(X, F) = 2\int \psi(y, F)\gamma(x, y, F)\, dF(y) + \psi^2(x, F) - \sigma^2(F).$$

Since

$$\tilde{T} = T\left(1 - \frac{\sigma^{-2}}{2}(F)\left(\hat{\sigma}^2(F) - \sigma^2(F)\right)\right) + O_P(n^{-1})$$

if we take K_{3n}, $\tilde{K}_{3n}$ to be the cumulants of the appropriate linear or quadratic approximations to T, $\tilde{T}$, we get

$$\tilde{K}_{3n} - K_{3n} = -3n^{-1/2}\Big[\mathrm{SKEW}(\psi) + \sigma^{-3}(F) \times \int\int \psi(x, F)\psi(y, F)\gamma(x, y, F)\, dF(x)\, dF(y)\Big] + O(n^{-1}) \tag{3}$$

and from Withers (1983),

$$K_{3n} = n^{-1/2}\Big[\mathrm{SKEW}(\psi) + \tfrac{3}{2}\sigma^{-3}(F) \times \int \psi(x)\psi(y)\gamma(x, y)\, dF(x)\, dF(y)\Big] + O(n^{-1}). \tag{4}$$

From (3) and (4) we get

$$K_{3n} + \tilde{K}_{3n} = -\mathrm{SKEW}(\psi) \tag{5}$$

and hence

$$a = \frac{\mathrm{SKEW}}{6}(\psi). \tag{6}$$

These computations and Efron [(1987), Section 6] suggest that if **P** is any model (parametric, fully nonparametric or semiparametric), the map $F \to \theta(F)$ is sufficiently smooth, $\hat{\theta}$ is an efficient (on **P**) estimate of $\theta(F)$ and $\hat{\sigma}$ is an efficient (on **P**) estimate of the asymptotic variance of $n^{1/2}\hat{\theta}$, then (6) should hold. Of course, to make any such assertions precise we need to assume existence of stochastic (von Mises) and Edgeworth expansions for T and $\tilde{T}$. For appropriate conditions and the necessary formulae, see Pfanzagl [(1985), Chapter 10], where an extensive discussion of questions such as those treated in Section 4 may also be found.

REFERENCES

BICKEL, P. J. (1974). Edgeworth expansions in nonparametric statistics. *Ann. Statist.* **2** 1–20.

EFRON, B. (1987). Better bootstrap confidence intervals (with discussion). *J. Amer. Statist. Assoc.* **82** 171–200.

HUBER, P. J. (1981). *Robust Statistics*. Wiley, New York.

PFANZAGL, J. (1985). *Asymptotic Expansions for General Statistical Models. Lecture Notes in Statist.* **31**. Springer, Berlin.

WITHERS, C. S. (1983). Expansions for the distribution and quantiles of a regular functional of the empirical distribution with application to nonparametric confidence intervals. *Ann. Statist.* **11** 577–587.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720