Inconsistent Hybrid Bootstrap Confidence Regions

Robert W. Keener^{a,*}, Hokeun Sun^b

^a*Department of Statistics, University of Michigan, Ann Arbor, MI 48103, USA*

^b*Department of Statistics, Pusan National University, Pusan, South Korea*

Abstract

In a simple signal plus noise model for Poisson data, of interest in high energy physics, coverage probabilities for hybrid bootstrap confidence intervals do not converge to the desired nominal value as information about the nuisance parameter increases.

Keywords: Hybrid Bootstrap, confidence intervals, resampling, bootstrap

2000 MSC: 62F40, 62F25, 62E20

1. Introduction

A standard approach used to find a confidence set for an unknown parameter $\theta \in \Omega$ based on data $X \sim P_\theta$ is to invert the family of likelihood ratio tests. Specifically, if $l(\cdot)$ is the log likelihood function, if $\hat{\theta}$ is the maximum likelihood estimator, if

$$\Lambda(\theta_0) = l(\hat{\theta}) - l(\theta_0)$$

is the generalized log likelihood ratio test statistic used to test $\theta = \theta_0$ versus $\theta \neq \theta_0$, and if $q(\theta)$ is the upper α -th quantile for the P_θ distribution of $\Lambda(\theta)$,

*Corresponding author

Email addresses: `keener@umich.edu` (Robert W. Keener), `hsun@pusan.ac.kr` (Hokeun Sun)

then

$$\{\theta : \Lambda(\theta) \leq q(\theta)\}$$

is a $1 - \alpha$ confidence region for θ .

The hybrid bootstrap, introduced to set confidence intervals in group sequential tests in Chuang and Lai (1998), is considered in a general context in Chuang and Lai (2000). The method extends the duality approach just described to interval estimation for a parameter of interest when there are nuisance parameters. Let θ and η denote the parameter of interest and the nuisance parameter, and let $\hat{\theta}$ and $\hat{\eta}$ be the maximum likelihood estimators for these parameters. If $\hat{\eta}_\theta$ maximizes the log likelihood $l(\theta, \eta)$ over η with θ fixed, then the log likelihood test statistic to test $\theta = \theta_0$ versus $\theta \neq \theta_0$ is now

$$\Lambda(\theta_0) = l(\hat{\theta}, \hat{\eta}) - l(\theta_0, \hat{\eta}_{\theta_0}).$$

Let $q(\theta, \eta)$ denote the upper α -th quantile for $\Lambda(\theta)$ under $P_{\theta, \eta}$. The region

$$\{\theta : \Lambda(\theta) \leq q(\theta, \eta), \forall \eta\} \tag{1}$$

has coverage probability at least $1 - \alpha$. But to find it, quantiles $q(\theta, \eta)$ are needed for all θ and η , and the region may be too conservative. The ordinary bootstrap confidence region is

$$\{\theta : \Lambda(\theta) \leq q(\hat{\theta}, \hat{\eta})\}. \tag{2}$$

The only quantile necessary to compute this region is $q(\hat{\theta}, \hat{\eta})$. This quantile can be found (if necessary) by bootstrap simulation generating data X^* from $P_{\hat{\theta}, \hat{\eta}}$. The $P_{\theta, \eta}$ coverage of this interval will be approximately $1 - \alpha$ if $q(\hat{\theta}, \hat{\eta})$ accurately estimates $q(\theta, \eta)$. In regular models with large samples this will

be the case for two reasons: the maximum likelihood estimators $\hat{\theta}$ and $\hat{\eta}$ are consistent, and the null distributions for $\Lambda(\theta)$ are approximately independent of θ . In practice, the bootstrap region (2) often works well with moderate sample sizes, but with smaller samples its performance is suspect.

The hybrid bootstrap confidence region is

$$S = S(X) = \{\theta : \Lambda(\theta) \leq q(\theta, \hat{\eta}_\theta)\}. \quad (3)$$

To compute S , quantiles $q(\theta, \hat{\eta}_\theta)$ are necessary for all θ . These can be found by bootstrap simulation generating data X_θ^* from $P_{\theta, \hat{\eta}_\theta}$ for values of θ in a reasonably fine grid. Since multiple simulations are required, the computational burden to compute the hybrid region S is greater than that for the ordinary bootstrap, especially if η is multidimensional. But with modern computing the computations are often feasible. Note that bootstrap simulations to find $q(\theta, \eta)$ for all θ and η to compute the first interval (1) would need to be done for a grid of values for θ and η , posing a greater burden than the simulations necessary for the hybrid region S in (3).

The $P_{\theta, \eta}$ coverage for the hybrid region S will be approximately $1 - \alpha$ if $q(\theta, \hat{\eta}_\theta)$ accurately estimates $q(\theta, \eta)$. As with the bootstrap region, this should be the case in large samples but may be suspect with small samples. But there are several interesting examples in which the data provide substantial information about the nuisance parameter η , but limited information about the parameter of interest θ . In these cases, the hybrid region may perform much better than the ordinary bootstrap region.

The next section considers a simple signal plus noise model for Poisson data, of interest in high energy physics. In this example coverage probabilities for the hybrid region S do not converge to the desired nominal value as

information about the nuisance parameter increases. Inconsistency arises because test statistics for θ are discrete when η is known.

When the nuisance parameter η is known, exact coverage can be achieved in a standard fashion by introducing extraneous randomization. Although this gives smaller intervals and increases the sensitivity of associated tests, randomization feels artificial and this approach may not have much appeal. The situation is similar when η is unknown. If desired, estimation error for $\hat{\eta}$ can be used to construct variables that are approximately uniformly distributed on $(0, 1)$; and these variables can be used to construct consistent confidence intervals analogous to the exact intervals based on external randomization when the η is known. But this approach uses $\hat{\eta}$ in a rather discontinuous fashion and feels artificial. Unfortunately, intervals that use the data in a more natural fashion will be inconsistent. Similar difficulties will almost certainly arise in other problems that are discrete when the nuisance parameter is known.

2. Poisson Example

Researchers in high energy physics are at times interested in estimating a rate $\theta \geq 0$ from a Poisson measurement X with mean $\theta + \eta$. Here η represents a background rate, often considered as known from prior or “off-line” experiments. Also, in many cases $\theta = 0$ is a definite possibility, corresponding to the absence of the particle or phenomena the experiment is trying to detect. This problem is a bit nonstandard since EX is known to be at least η , and there has been some discussion in the physics literature about the proper way to set a confidence interval for θ . The “unified method” of Feldman and

Cousins (1998) amounts to inverting the family of likelihood ratio tests, and has seen wide interest in physics since its appearance. Related alternatives are discussed in Mandelkern (2002), Roe and Woodroffe (1999), and Roe and Woodroffe (2000).

In practice, the assumption that the background rate η is known is too optimistic. More realistically, information about η may come from count data Y modeled as Poisson with mean $\gamma\eta$. Here the scale factor γ , represents the ratio of the observation times for Y and X . With large γ there is considerable information about the background η , exactly the setting in which the hybrid bootstrap approach seems most promising.

Sen et al. (2009) have investigated the performance of the hybrid bootstrap confidence interval S in this example, and in their numerical work it seems to perform well. Feldman (2000) also considers this problem, although it is not clear if his suggestions are exactly the same as the hybrid bootstrap.

Despite the positive accounts of the hybrid bootstrap interval's performance, in this example S is not consistent: as $\gamma \rightarrow \infty$, $P_{\theta,\eta}(\theta \in S) \not\rightarrow 1 - \alpha$. The problem has to do with discreteness. Note that if the background η were known, exact coverage based on data X would be impossible without external randomization.

For notation, let

$$l_0(\theta, \eta) = X \log(\theta + \eta) - \theta - \eta - \log(X!),$$

the log likelihood from X , and

$$l_\gamma(\eta) = Y \log(\gamma\eta) - \gamma\eta - \log(Y!),$$

the log likelihood from Y . Then the total log likelihood is

$$l(\theta, \eta) = l_0(\theta, \eta) + l_\gamma(\eta).$$

Taking $\tilde{\eta} = Y/\gamma$, the maximum likelihood estimator for η based on Y , the maximum likelihood estimator for θ is

$$\hat{\theta} = (X - \tilde{\eta})_+,$$

and the maximum likelihood estimator for η is

$$\hat{\eta} = \begin{cases} \tilde{\eta}, & X \geq \tilde{\eta}; \\ \frac{X+Y}{1+\gamma}, & X \leq \tilde{\eta}. \end{cases}$$

Finally, let $\hat{\theta}_\eta = (X - \eta)_+$, the maximum likelihood estimator for θ when η is known, and let $\hat{\eta}_\theta$ denote the maximum likelihood estimator for η when θ is known, given explicitly as

$$\begin{aligned} \hat{\eta}_\theta &= \frac{X + Y - \theta(1 + \gamma) + \sqrt{[X + Y - \theta(1 + \gamma)]^2 + 4\theta Y(1 + \gamma)}}{2(1 + \gamma)}, \\ &= \frac{1}{2}(\tilde{\eta} - \theta) + \frac{X - \tilde{\eta}}{2(1 + \gamma)} \\ &\quad + \frac{1}{2}\sqrt{(\tilde{\eta} + \theta)^2 + \frac{2(\tilde{\eta} - \theta)X - 2\tilde{\eta}(\tilde{\eta} + \theta)}{1 + \gamma} + \frac{(X - \tilde{\eta})^2}{(1 + \gamma)^2}}. \end{aligned}$$

From the second equation here and the equation above for $\hat{\eta}$, it is easy to see that

$$\hat{\eta}_\theta = \tilde{\eta} + O_p(1/\gamma) \quad \text{and} \quad \hat{\eta} = \tilde{\eta} + O_p(1/\gamma) \quad (4)$$

as $\gamma \rightarrow \infty$. By Taylor expansion about $\tilde{\eta}$, with η_* an intermediate value between $\hat{\eta}$ and $\tilde{\eta}$,

$$l_\gamma(\hat{\eta}) = l_\gamma(\tilde{\eta}) - \frac{\gamma\tilde{\eta}(\hat{\eta} - \tilde{\eta})^2}{2\eta_*^2} = l_\gamma(\tilde{\eta}) + O_p(1/\gamma). \quad (5)$$

Similarly,

$$l_\gamma(\hat{\eta}_\theta) = l_\gamma(\tilde{\eta}) + O_p(1/\gamma). \quad (6)$$

Since $\hat{\eta}_\theta \xrightarrow{p} \eta$ as $\gamma \rightarrow \infty$, using (5) and (6)

$$\begin{aligned} \Lambda(\theta) &\xrightarrow{p} \Lambda_0(\theta, \eta) \\ &\stackrel{\text{def}}{=} l_0(\hat{\theta}_\eta, \eta) - l_0(\theta, \eta) \\ &= X \log \left(\frac{(X - \eta)_+ + \eta}{\theta + \eta} \right) + \theta - (X - \eta)_+, \end{aligned}$$

the log likelihood ratio test statistic when the background η is known. Let $q_0(\theta, \eta)$ denote the upper α -th quantile for $\Lambda_0(\theta, \eta)$. Since $\Lambda_0(\theta, \eta)$ is a function of X , it is a discrete variable, and to be precise and avoid ambiguity we will take

$$q_0(\theta, \eta) = \inf \{x : P_{\theta, \eta}(\Lambda_0(\theta, \eta) \leq x) \geq 1 - \alpha\}.$$

(The quantile $q(\theta, \eta)$ for $\Lambda(\theta)$ should also be defined similarly, although this feels less important since probabilities for the atoms of the distribution of $\Lambda(\theta)$ are all small.) Then $q_0(\theta, \eta)$ will be an atom for the $P_{\theta, \eta}$ distribution of $\Lambda_0(\theta, \eta)$, and since cumulative distribution functions are right continuous,

$$P_{\theta, \eta}(\Lambda_0(\theta, \eta) \leq q_0(\theta, \eta)) \geq 1 - \alpha$$

and

$$P_{\theta, \eta}(\Lambda_0(\theta, \eta) < q_0(\theta, \eta)) < 1 - \alpha.$$

Define

$$\Delta_+(\theta, \eta) = P_{\theta, \eta}(\Lambda_0(\theta, \eta) \leq q_0(\theta, \eta)) - (1 - \alpha) \geq 0,$$

$$\Delta_-(\theta, \eta) = 1 - \alpha - P_{\theta, \eta}(\Lambda_0(\theta, \eta) < q_0(\theta, \eta)) > 0,$$

and

$$\Delta(\theta, \eta) = \Delta_+(\theta, \eta) + \Delta_-(\theta, \eta) = P_{\theta, \eta}(\Lambda_0(\theta, \eta) = q_0(\theta, \eta)).$$

If $\Delta_+(\theta, \eta) > 0$, it is fairly easy to argue that $q(\theta, \eta)$ will converge to $q_0(\theta, \eta)$ as $\gamma \rightarrow \infty$. But a more careful analysis, based on the following technical lemma, is necessary to approximate probabilities.

Lemma 2.1. *Let W_γ and Z_γ be random variables with*

$$W_\gamma \xrightarrow{p} W \quad \text{and} \quad Z_\gamma \Rightarrow Z$$

as $\gamma \rightarrow \infty$. If W is a discrete variable with an atom at q , if Z_γ and W are asymptotically independent, and if

$$\sqrt{\gamma}(W_\gamma - W) = Z_\gamma \quad \text{on} \quad W = q,$$

then

$$P(W_\gamma \leq q + c/\sqrt{\gamma}) \rightarrow P(W < q) + P(W = q)P(Z \leq c)$$

whenever $P(Z = c) = 0$.

Proof. If $w < q$, since $W_\gamma \xrightarrow{p} W$

$$P(W_\gamma \leq q + c/\sqrt{\gamma}, W = w) \rightarrow P(W = w).$$

Summing over $w < q$, by dominated convergence

$$P(W_\gamma \leq q + c/\sqrt{\gamma}, W < w) \rightarrow P(W < w).$$

Similarly,

$$P(W_\gamma \leq q + c/\sqrt{\gamma}, W > w) \rightarrow 0.$$

Finally, if $P(Z = c) = 0$, by the asymptotic independence

$$\begin{aligned} P(W_\gamma \leq q + c/\sqrt{\gamma}, W = q) &= P(Z_\gamma \leq c, W = q) \\ &\rightarrow P(W = q)P(Z \leq c), \end{aligned}$$

and the lemma follows. $\square$

The main results below require two regularity assumptions:

A1: $\Delta_+(\theta, \eta) > 0$.

A2: There is a unique constant x^* so that $\Lambda_0(\theta, \eta) = q_0(\theta, \eta)$ if and only if $X = x^*$.

A1 ensures that q_0 is continuous at (θ, η) , so that small changes for the values of these parameters will not have a large effect on this quantile. As a function of X , $\Lambda_0(\theta, \eta)$ is unimodal, achieving its minimum when $X = \theta + \eta$. So it is possible for two integral values for X to give the same value for $\Lambda_0(\theta, \eta)$, but the assertion in A2 should be typical. If θ and η are parameter values satisfying A2, introduce

$$\tilde{q}(u) = \tilde{q}_{\theta, \eta}(u) = x^* \log \left(\frac{\max\{x^*, u\}}{\theta + u} \right) + \theta - (x^* - u)_+.$$

Note that for u sufficiently close to η , $\tilde{q}(u)$ will be the value for $\Lambda_0(\theta, \eta)$ when $X = x^*$, and that $\tilde{q}$ is differentiable with

$$\tilde{q}'(\eta) = \frac{\min\{x^*, \eta\}}{\eta} - \frac{x^*}{\theta + \eta}.$$

Proposition 2.2. *Let θ and η be parameter values satisfying A1 and A2, define*

$$\sigma^2(\theta, \eta) = \eta [\tilde{q}'(\eta)]^2 = \eta \left[\frac{\min\{x^*, \eta\}}{\eta} - \frac{x^*}{\theta + \eta} \right]^2,$$

and assume that $\sigma(\theta, \eta) > 0$. Then for any $c \in \mathbb{R}$, as $\gamma \rightarrow \infty$,

$$\begin{aligned} P_{\theta, \eta}[\Lambda(\theta) \leq q_0(\theta, \eta) + c/\sqrt{\gamma} + o_p(1/\sqrt{\gamma})] \\ \rightarrow 1 - \alpha - \Delta_-(\theta, \eta) + \Delta(\theta, \eta)\Phi(c/\sigma(\theta, \eta)) \end{aligned}$$

and

$$q(\theta, \eta) = q_0(\theta, \eta) + \frac{\sigma(\theta, \eta)}{\sqrt{\gamma}} \Phi^{-1} \left(\frac{\Delta_-(\theta, \eta)}{\Delta(\theta, \eta)} \right) + o(1/\sqrt{\gamma}),$$

where Φ^{-1} is the inverse of the standard normal cumulative distribution function Φ . Also, the results here hold uniformly for η sufficiently close to a value satisfying A1 and A2.

Remark 2.3. The restriction $\sigma(\theta, \eta) > 0$ is fairly innocuous, but does fail if $x^* = 0$.

Proof. By (4),

$$\begin{aligned} l_0(\hat{\theta}, \hat{\eta}) - l_0(\theta, \hat{\eta}_\theta) &= l_0[(X - \tilde{\eta})_+, \tilde{\eta}] - l_0(\theta, \tilde{\eta}) + O_p(1/\gamma) \\ &= \Lambda_0(\theta, \tilde{\eta}) + O_p(1/\gamma), \end{aligned}$$

and using this, (5), and (6),

$$\Lambda(\theta) = \Lambda_0(\theta, \tilde{\eta}) + O_p(1/\gamma).$$

From this, $\Lambda(\theta) \xrightarrow{p} \Lambda_0(\theta, \eta)$ and we are in a setting where Lemma 2.1 is relevant. By A2, on $\Lambda_0(\theta, \eta) = q_0(\theta, \eta)$, $X = x^*$, $\Lambda_0(\theta, \tilde{\eta}) = \tilde{q}(\tilde{\eta})$, and $\Lambda_0(\theta, \eta) = \tilde{q}(\eta)$. By the delta method, on $X = x^*$,

$$\begin{aligned} \sqrt{\gamma}[\Lambda(\theta) - \Lambda_0(\theta, \eta)] &= \sqrt{\gamma}[\tilde{q}(\tilde{\eta}) - \tilde{q}(\eta)] + o_p(1) \\ &= \sqrt{\gamma}(\tilde{\eta} - \eta)\tilde{q}'(\eta) + o_p(1) \\ &\Rightarrow N\left(0, \eta[\tilde{q}'(\eta)]^2\right), \end{aligned} \tag{7}$$

by normal approximation for the Poisson distribution. The first assertion of the proposition now follows from Lemma 2.1. The second assertion of the lemma follows easily from the first. Uniformity in η is evident from the method of proof. $\square$

The coverage probability for the hybrid confidence region (3) is

$$P_{\theta,\eta}[\Lambda(\theta) \leq q(\theta, \hat{\eta}_\theta)]. \quad (8)$$

This can be approximated using Lemma 2.1, although there is a bit of degeneracy. By Proposition 2.2,

$$q(\theta, \hat{\eta}_\theta) = q_0(\theta, \hat{\eta}_\theta) + c/\sqrt{\gamma} + o_p(1/\sqrt{\gamma})$$

with

$$c = \sigma(\theta, \eta) \Phi^{-1} \left(\frac{\Delta_-(\theta, \eta)}{\Delta(\theta, \eta)} \right).$$

Now $q_0(\theta, \eta) = \tilde{q}(\eta)$, and since $\hat{\eta}_\theta \xrightarrow{p} \eta$, with probability tending to one

$$q_0(\theta, \hat{\eta}_\theta) = \tilde{q}(\hat{\eta}_\theta) = \tilde{q}(\tilde{\eta}) + o_p(1/\sqrt{\gamma}).$$

So the coverage probability (8) is

$$P_{\theta,\eta}[\Lambda(\theta) + \tilde{q}(\eta) - \tilde{q}(\tilde{\eta}) + o_p(1/\sqrt{\gamma}) \leq q_0(\theta, \eta) + c/\sqrt{\gamma}].$$

The variable on the left hand side of the inequality here converges in probability to $\Lambda_0(\theta, \eta)$. But by (7), on $X = x^*$,

$$\sqrt{\gamma}[\Lambda(\theta) + \tilde{q}(\eta) - \tilde{q}(\tilde{\eta}) - \Lambda_0(\theta, \eta)] \xrightarrow{p} 0,$$

and so we have weak convergence to a degenerate distribution. Lemma 2.1 then gives the following corollary which shows that the hybrid confidence interval is generally inconsistent.

Corollary 2.4. *If A1 and A2 hold and $\Delta_-(\theta, \eta) \neq \Delta(\theta, \eta)/2$, the coverage probability for the hybrid bootstrap confidence region (3) converges to*

$$1 - \alpha - \Delta_-(\theta, \eta) + \Delta(\theta, \eta)I\{\Delta_-(\theta, \eta) > \Delta(\theta, \eta)/2\}$$

as $\gamma \rightarrow \infty$.

For consistent coverage, quantiles for $\Lambda(\theta)$ would have to be estimated with error $o_p(1/\sqrt{\gamma})$, and to achieve this, η would need to be estimated with the same precision, which is not possible. The ordinary bootstrap is known to break down if the functionals of interest are not sufficiently smooth, and this corollary may be seen as an analogous result for the hybrid bootstrap.

Example 2.5. If $\theta = 1/2$ and $\alpha = 10\%$, then, for $\eta \in (.03182, .60207)$ we have $x^* = 2$. Coverage probabilities in this case, computed numerically by summing against Poisson mass functions, are given in Table 1 for selected values for η and γ . In most cases, convergence to the limit in Corollary 2.4 is rapid. The exception is when $\eta = 0.2$. We suspect that slow convergence in this case occurs because $\Delta_- = 5.580\%$ is very close to $\Delta/2 = 6.083\%$.

Noting the inconsistent coverage in Corollary 2.4, one might hope that a suitable modification of the hybrid bootstrap method would give consistent confidence regions. Unfortunately, the difficulty here seems to be somewhat intrinsic. Let R denote any confidence region for θ that includes the parameter value θ according to the value of the test statistic $\Lambda(\theta)$, at least with high probability. Specifically, assume that as $\gamma \rightarrow \infty$,

$$I\{\theta \in R\} = I\{\Lambda(\theta) \leq c_\gamma\} + o_p(1), \quad (9)$$

or equivalently that the probability of the symmetric difference of these events tends to zero. Suppose $P_{\theta, \eta_0}(\theta \in R) \rightarrow 1 - \alpha$, with the parameter values θ

Table 1: Coverage Probabilities with $\theta = 1/2$

η	$\gamma = 10$	$\gamma = 50$	$\gamma = 100$	$\gamma = 250$	$\gamma = \infty$
0.1	0.9042	0.8812	0.8783	0.8781	0.8781
0.2	0.9165	0.9101	0.8978	0.8873	0.8442
0.3	0.9244	0.9425	0.9475	0.9521	0.9526
0.4	0.9246	0.9363	0.9371	0.9371	0.9371
0.5	0.9204	0.9197	0.9197	0.9197	0.9197
0.6	0.9159	0.9020	0.9005	0.9004	0.9004

and η_0 satisfying A1 and A2, and assume that $\tilde{q}'(\eta_0) \neq 0$. Then $P_{\theta, \eta_0}(\Lambda(\theta) \leq c_\gamma) \rightarrow 1 - \alpha$, and, using essentially the same argument as that used to prove Proposition 2.2,

$$c_\gamma = q_0(\theta, \eta_0) + \frac{\sigma(\theta, \eta_0)}{\sqrt{\gamma}} \Phi^{-1} \left(\frac{\Delta_-(\theta, \eta_0)}{\Delta(\theta, \eta_0)} \right) + o(1/\sqrt{\gamma}).$$

By A1, $q_0(\theta, \cdot)$ is continuous at η_0 . By A2, $\Lambda_0(\theta, \eta_0) = q_0(\theta, \eta)$ if and only if $X = x^*$. Then, since $\Lambda_0(\theta, \eta)$ is a continuous in η , for η in some neighborhood of η_0 we will also have

$$\Lambda_0(\theta, \eta) = q_0(\theta, \eta) \quad \text{if and only if} \quad X = x^*$$

and

$$\Lambda_0(\theta, \eta) < q_0(\theta, \eta) \quad \text{if and only if} \quad \Lambda_0(\theta, \eta_0) < q_0(\theta, \eta_0).$$

Considering probabilities of these events, since the $P_{\theta, \eta}$ distributions for X converge strongly to the P_{θ, η_0} distributions for X as $\eta \rightarrow \eta_0$, for η in a

possibly smaller neighborhood of η_0 , we will also have

$$P_{\theta,\eta}[\Lambda_0(\theta, \eta) < q_0(\theta, \eta)] < 1 - \alpha.$$

Since $\tilde{q}'(\eta_0) \neq 0$, in this neighborhood we can find a value η with $q_0(\theta, \eta) > q_0(\theta, \eta_0)$. But then,

$$\begin{aligned} P_{\theta,\eta}[\Lambda(\theta) \leq c_\gamma, \Lambda_0(\theta, \eta) \geq q_0(\theta, \eta)] \\ \leq P_{\theta,\eta}[|\Lambda(\theta) - \Lambda_0(\theta, \eta)| \geq q_0(\theta, \eta) - c_\gamma], \end{aligned}$$

which tends to zero since $q_0(\theta, \eta) - c_\gamma \rightarrow q_0(\theta, \eta) - q_0(\theta, \eta_0) > 0$ and $\Lambda(\theta) \xrightarrow{p} \Lambda_0(\theta, \eta)$. So

$$\begin{aligned} \limsup P_{\theta,\eta}(\theta \in R) &= \limsup P_{\theta,\eta}[\Lambda(\theta) \leq c_\gamma] \\ &\leq P_{\theta,\eta}[\Lambda_0(\theta, \eta) < q_0(\theta, \eta)] < 1 - \alpha, \end{aligned}$$

showing that coverage probabilities for R cannot converge to $1 - \alpha$.

Requirement (9) might be viewed as insisting that the information from Y (or $\tilde{\eta}$) be used in a sensible fashion. Without this restriction, consistent confidence regions are possible. One simple way to do this is to base tests on

$$\tilde{X} = X + U_\gamma$$

with

$$U_\gamma = \sqrt{\gamma}\tilde{\eta} - \lfloor \sqrt{\gamma}\tilde{\eta} \rfloor,$$

the fractional part of $\sqrt{\gamma}\tilde{\eta} = Y/\sqrt{\gamma}$. If U is uniformly distributed on $(0, 1)$ and independent of X , then

$$U_\gamma \Rightarrow U \quad \text{and} \quad \tilde{X} \Rightarrow X + U$$

as $\gamma \rightarrow \infty$. If $q^+(\theta, \eta)$ and $q^-(\theta, \eta)$ are upper and lower $\alpha/2$ -th quantiles for $\tilde{X}$, and if $q_0^+(\theta, \eta)$ and $q_0^-(\theta, \eta)$ are the corresponding upper and lower $\alpha/2$ -th quantiles for $X + U$, then it is not hard to show that

$$q^\pm(\theta, \hat{\eta}_\theta) \xrightarrow{p} q_0^\pm(\theta, \eta)$$

and

$$P_{\theta, \eta}[q^-(\theta, \hat{\eta}_\theta) < \tilde{X} < q^+(\theta, \hat{\eta}_\theta)] \rightarrow 1 - \alpha$$

as $\gamma \rightarrow \infty$. It follows that coverage probabilities for the the hybrid confidence region

$$\{\theta : q^-(\theta, \hat{\eta}_\theta) < \tilde{X} < q^+(\theta, \hat{\eta}_\theta)\}$$

converge to $1 - \alpha$ as $\gamma \rightarrow \infty$.

Confidence regions might also be based on a modified version of $\Lambda(\theta)$, obtained increasing it by a suitable multiple of U_γ . But if the multiple is

$$\delta(\theta) = \inf_{x: \Lambda(\theta; x, \tilde{\eta}) > \Lambda(\theta)} \Lambda(\theta; x, \tilde{\eta}) - \Lambda(\theta),$$

so the modified statistic is $\Lambda(\theta) + \delta(\theta)U_\gamma$, it seems that consistency may fail if A2 does not hold. In testing problems with discrete data, external randomization is generally necessary to achieve arbitrary significance levels. This region and the region based on $\tilde{X}$ achieve the necessary randomization asymptotically using U_γ . These regions are rather ad hoc and would be more appealing if they used Y in a more natural fashion.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. 0706771. Any opinions, findings, and conclusions or

recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

- Chuang, C. and T. L. Lai (1998). Resampling methods for confidence intervals in group sequential trials. *Biometrika* 85, 317–332.
- Chuang, C. and T. L. Lai (2000). Hybrid resampling methods for confidence intervals. *Statist. Sinica* 10, 1–50.
- Feldman, G. (2000). Multiple measurements and parameters in the unified approach. Talk at the FermiLab Workshop on Confidence Limits. <http://www.hepl.harvard.edu/~feldman/CL2k.pdf>.
- Feldman, G. and R. Cousins (1998). Unified approach to the classical statistical analysis of small signal. *Phys. Rev. D* 57, (38733889).
- Mandelkern, M. (2002). Setting confidence intervals for bounded parameters. *Statist. Sci.* 17, 149–172.
- Roe, B. and M. Woodroffe (1999). Improved probability method for estimating signal in the presence of background. *Phys. Rev. D* 60, (053009).
- Roe, B. and M. Woodroffe (2000). Setting confidence belts. *Phys. Rev. D.* 63, (013009).
- Sen, B., M. Walker, and M. Woodroffe (2009). On the unified method with nuisance parameters. *Statist. Sinica* 19, 301–314.