Fiducial Generalized Confidence Intervals

Jan HANNIG, Hari IYER, and Paul PATTERSON

Generalized pivotal quantities (GPQs) and generalized confidence intervals (GCIs) have proven to be useful tools for making inferences in many practical problems. Although GCIs are not guaranteed to have exact frequentist coverage, a number of published and unpublished simulation studies suggest that the coverage probabilities of such intervals are sufficiently close to their nominal value so as to be useful in practice. In this article we single out a subclass of generalized pivotal quantities, which we call *fiducial generalized pivotal quantities* (FGPQs), and show that under some mild conditions, GCIs constructed using FGPQs have correct frequentist coverage, at least asymptotically. We describe three general approaches for constructing FGPQs—a recipe based on invertible pivotal relationships, and two extensions of it—and demonstrate their usefulness by deriving some previously unknown GPQs and GCIs. It is fair to say that nearly every published GCI can be obtained using one of these recipes. As an interesting byproduct of our investigations, we note that the subfamily of fiducial generalized pivots has a close connection with fiducial inference proposed by R. A. Fisher. This is why we refer to the proposed generalized pivots as fiducial generalized pivotal quantities. We demonstrate these concepts using several examples.

KEY WORDS: Asymptotic properties; Common mean problem; Conditional inference; Fiducial inference; Generalized pivot; Structural inference; Structural method.

1. INTRODUCTION

Tsui and Weerahandi (1989) introduced the concept of generalized p values and generalized test variables, which are useful for developing hypothesis tests in situations where traditional frequentist approaches do not provide useful solutions. Subsequently, Weerahandi (1993) introduced the concept of a generalized pivotal quantity (GPQ) for a scalar parameter θ, which can be used to construct an interval estimator for θ in situations where standard pivotal quantity-based approaches may not be applicable. He referred to such intervals as generalized confidence intervals (GCIs). Since then, several GCIs have been constructed in many practical problems (see, e.g., Weerahandi 1995; Chang and Huang 2000; Hamada and Weerahandi 2000; McNally, Iyer, and Mathew 2001; Burdick and Park 2003; Krishnamoorthy and Lu 2003; Krishnamoorthy and Mathew 2003; Iyer, Wang, and Mathew 2004; Mathew and Krishnamoorthy 2004; Weerahandi 2004; Burdick, Borror, and Montgomery 2005; Burdick, Park, Montgomery, and Borror 2005; Daniels, Burdick, and Quiroz 2005; Roy and Matthew 2005). These intervals do not always have exact frequentist coverage. Nevertheless, results of simulation studies reported in the literature appear to support the claim that coverage probabilities of GCIs are sufficiently close to their stated value so that they are in fact useful procedures in practical problems. Despite the large number of successful applications of GCIs reported in the literature, it is surprising that there are no published theoretical results discussing either small-sample properties or asymptotic behavior of GCIs.

A simple test case for the application of GCIs is the Gaussian two-sample problem with heterogeneous variances, where one is interested in a confidence interval for the difference $\mu_1 - \mu_2$ between the two means. This is the well-known Behrens–Fisher problem for which Behrens (1929) proposed a solution and Fisher (1935) gave a justification based on the *fiducial argument*. Weerahandi (1993) derived a GPQ for $\mu_1 - \mu_2$ and remarked that the resulting interval coincided with the fiducial solution.

In this article we identify an important subclass of GPQs, which we call *fiducial generalized pivotal quantities* (FGPQs) for reasons that we discuss shortly. We also provide some general methods for constructing FGPQs for large classes of problems. Nearly every published GCI can be obtained using these methods. More important, and perhaps of greater interest to practitioners, we also show that, under reasonable assumptions, GCIs based on FGPQs have asymptotically correct frequentist coverage. This result provides a frequentist justification for GCIs (and also for generalized tests, although our focus here is confidence intervals) when the GPQ is chosen appropriately. In addition, we provide a number of examples to illustrate these results.

The reason that we chose the term "FGPQ" is that GCIs based on FGPQs are in fact obtainable using the fiducial argument of Fisher (1935) within a suitably chosen framework, such as the pivotal quantity approach of Barnard (1977, 1981, 1982, 1995), the structural inference of Fraser (1966, 1968), and the functional model basis for fiducial inference discussed by Dawid and Stone (1982). We establish the connection between FGPQs and fiducial distributions by showing that, given a fiducial distribution for a parameter, there is a systematic procedure for constructing a FGPQ whose distribution is the same as the fiducial distribution. As a byproduct, this connection of FGPQs with fiducial inference leads to a frequentist justification for fiducial inference in many settings. In fact, FGPQs provide a natural framework for associating a distribution with a parameter.

The article is organized as follows. In the next section we give a brief introduction to GPQs and GCIs and also introduce the subclass of FGPQs. In Section 3 we prove a theorem to the effect that under fairly general conditions, GCIs obtained from FGPQs have correct asymptotic coverage. Some familiar examples are considered, and the application of the theorem is illustrated. We also discuss an example where the conditions of the theorem are not satisfied.

Jan Hannig is Assistant Professor, Department of Statistics, Colorado State University, Fort Collins, CO 80523 (E-mail: *hannig@stat.colostate.edu*). Hari Iyer is Professor, Department of Statistics, Colorado State University, Fort Collins, CO 80523 (E-mail: *hari@stat.colostate.edu*). Paul Patterson is Graduate Student, Department of Statistics, Colorado State University, Fort Collins, CO 80523, and Mathematical Statistician, USFS-Rocky Mountain Research Station (E-mail: *PLPatterson@fs.fed.us*). The authors thank the reviewers and the associate editor for their constructive comments and valuable suggestions. Hannig's research was supported in part by National Science Foundation grant DMS-05-04737, and Iyer's research was partially supported by National Park Service contract H2380040002 T.O. 04-51.

Journal of the American Statistical Association
March 2006, Vol. 101, No. 473, Theory and Methods
DOI 10.1198/016214505000000736

Section 4 is devoted to some general methods for constructing FGPQs. First, we describe a recipe for constructing FGPQs and discuss the scope of application of this recipe. This is a reformulation, using the notation of this article, of the recipe for constructing FGPQs given by Iyer and Patterson (2002). The procedure is illustrated with some examples of GPQs not previously discussed in the literature. We show that these GPQs (actually FGPQs) satisfy the conditions of the main theorem, so that the resulting GCIs are guaranteed to have the correct frequentist coverage asymptotically. In Section 5 we introduce two additional methods for constructing FGPQs that extend the range of problems for which GCIs can be developed. We illustrate the application of these methods with new confidence intervals for some well-known problems.

In Section 6 we discuss connections between GPQs and fiducial inference. We also touch on nonuniqueness issues associated with GCIs and fiducial intervals. We provide some concluding remarks in Section 7.

2. GENERALIZED PIVOTAL QUANTITIES AND GENERALIZED CONFIDENCE INTERVALS

2.1 Generalized Pivotal Quantities

The definition that we present for a GPQs is superficially different from Weerahandi's (1993) definition but is identical in spirit. Our definition enables us to more clearly explain the connection between ordinary pivotal quantities and GPQs and also facilitates a mathematically rigorous discussion of the asymptotic behavior of the resulting GCIs.

Definition 1. Let $\mathbb{S} \in \mathbb{R}^k$ denote an observable random vector whose distribution is indexed by a (possibly vector) parameter $\xi \in \mathbb{R}^p$. Suppose that we are interested in making inferences about $\theta = \pi(\xi) \in \mathbb{R}^q$ $(q \geq 1)$. Let $\mathbb{S}^\star$ represent an independent copy of $\mathbb{S}$. We use $\mathbf{s}$ and $\mathbf{s}^\star$ to denote realized values of $\mathbb{S}$ and $\mathbb{S}^\star$. A GPQ for θ, denoted by $\mathcal{R}_\theta(\mathbb{S}, \mathbb{S}^\star, \xi)$ (or simply $\mathcal{R}_\theta$ or $\mathcal{R}$ when there is no ambiguity) is a function of $(\mathbb{S}, \mathbb{S}^\star, \xi)$ with the following properties:

(GPQ1) The conditional distribution of $\mathcal{R}_\theta(\mathbb{S}, \mathbb{S}^\star, \xi)$, conditional on $\mathbb{S} = \mathbf{s}$, is free of ξ.
(GPQ2) For every allowable $\mathbf{s} \in \mathbb{R}^k$, $\mathcal{R}_\theta(\mathbf{s}, \mathbf{s}, \xi)$ depends on ξ only through θ.

Note that (GPQ2) uses $\mathbf{s}$ in both the first and the second argument positions of $\mathcal{R}_\theta(\cdot, \cdot, \xi)$. This is an important aspect of the definition of GPQ and explains both its similarity to and difference from an ordinary pivotal quantity. $\mathcal{R}_\theta(\mathbb{S}, \mathbb{S}^\star, \xi)$ would indeed be an ordinary pivotal quantity based on $(\mathbb{S}, \mathbb{S}^\star)$ if in condition (GPQ2) we instead had the requirement that $\mathcal{R}_\theta(\mathbf{s}, \mathbf{s}^\star, \xi)$ depend on ξ only through θ. The point is that we do not intend to actually observe a realization of $\mathbb{S}^\star$ and thus are not seeking a pivotal quantity for θ based on both $\mathbb{S}$ and $\mathbb{S}^\star$. However, heuristic reasoning suggests that under appropriate conditions, $\mathbb{S}$ and $\mathbb{S}^\star$ will be sufficiently close to each other, and hence we can substitute $\mathbb{S}$ in both argument positions in $\mathcal{R}_\theta(\cdot, \cdot, \xi)$ and use the distribution from condition (GPQ1) to make approximate confidence statements about the quantity $\mathcal{R}_\theta(\mathbf{s}, \mathbf{s}, \xi)$, and hence about θ. Theorem 1 in Section 3 confirms that our heuristic reasoning is indeed valid.

Property (GPQ2) implies that $\mathcal{R}_\theta(\mathbf{s}, \mathbf{s}, \xi) = f(\mathbf{s}, \theta)$ for some function f. It turns out that the subclass of GPQs for which $f(\mathbf{s}, \theta)$ is a function of θ only, say $f(\mathbf{s}, \theta) = f(\theta)$, have a special connection with fiducial inference. Generalized confidence regions obtained using such GPQs are not guaranteed to be intervals unless the function $f(\theta)$ is invertible. In this case one may assume that $f(\theta)$ is identically equal to θ without loss in generality. Such GPQs exist in practically every application that we have considered. This leads us to single out the following subclass of GPQs.

Definition 2. A GPQ $\mathcal{R}_\theta(\mathbb{S}, \mathbb{S}^\star, \xi)$ for a parameter θ is called a *fiducial generalized pivotal quantity* (FGPQ) if it satisfies the following two conditions:

(FGPQ1) The conditional distribution of $\mathcal{R}_\theta(\mathbb{S}, \mathbb{S}^\star, \xi)$, conditional on $\mathbb{S} = \mathbf{s}$, is free of ξ.
(FGPQ2) For every allowable $\mathbf{s} \in \mathbb{R}^k$, $\mathcal{R}_\theta(\mathbf{s}, \mathbf{s}, \xi) = \theta$.

Remark 1. Condition (FGPQ1) of Definition 2 is the same as condition (GPQ1) of Definition 1, but condition (FGPQ2) is a stronger version of condition (GPQ2) in the definition of a GPQ.

In Section 6 we show that if $\mathcal{R}_\theta$ is a FGPQ for θ, then frequentist probability intervals associated with the distribution of $\mathcal{R}_\theta$ have a corresponding interpretation as fiducial probability intervals associated with the parameter θ. It is for this reason that we use the term FGPQ to describe members of this important subclass.

2.2 Some Well-Known Generalized Pivotal Quantities

Example 1: The Behrens–Fisher problem. Consider m iid observations $X_i, i = 1, \dots, m$, from $\mathrm{N}(\mu_X, \sigma_X^2)$ and n iid observations $Y_j, j = 1, \dots, n$, from $\mathrm{N}(\mu_Y, \sigma_Y^2)$, where μ_X, μ_Y, σ_X, and σ_Y are unknown parameters. The problem is to obtain confidence bounds for the difference $\theta = \mu_X - \mu_Y$. Let $\bar{X}$ and $\bar{Y}$ denote the sample means and let S_X^2 and S_Y^2 denote the sample variances for the two samples. Then $\bar{X} \sim \mathrm{N}(\mu_X, \sigma_X^2/m)$, $\bar{Y} \sim \mathrm{N}(\mu_Y, \sigma_Y^2/n)$, $(m-1)S_X^2/\sigma_X^2 \sim \chi^2(m-1)$, and $(n-1)S_Y^2/\sigma_Y^2 \sim \chi^2(n-1)$. The statistic $\mathbb{S} = (\bar{X}, \bar{Y}, S_X^2, S_Y^2)$ is complete and sufficient for $\xi = (\mu_X, \mu_Y, \sigma_X, \sigma_Y)$.

A nontrivial exact confidence interval, in the frequentist sense, is unavailable for this problem (see, e.g., Linnik 1968). A solution to this problem was put forward by Behrens (1929); later, Fisher (1935) showed that this solution could be derived very simply using the fiducial argument. The fiducial distribution of $\mu_X - \mu_Y$ is known as the Behrens–Fisher distribution. Critical values for this distribution have been tabulated by Sukhatme (1958). The quantiles of this distribution lead to the lower and upper confidence bounds for $\mu_X - \mu_Y$.

Many other approximate confidence interval procedures for this problem have been discussed in the literature (see, e.g., Welch 1947; Satterthwaite 1942, 1946; Cochran 1964; Graybill and Wang 1980). Weerahandi (1993) derived a GCI for $\mu_X - \mu_Y$ and noted that it is the same as the Behrens–Fisher solution. He derived the GCI for $\theta = \mu_X - \mu_Y$ by starting with the statistics $\bar{X} - \bar{Y}$, S_X^2, and S_Y^2. He justified this based on invariance arguments. The GPQ proposed by Weerahandi is

$$\mathcal{R} = \mathcal{R}(\mathbb{S}, \mathbb{S}^\star, \xi)$$
$$= (\bar{X}^\star - \bar{Y}^\star - \theta)\left(\frac{\sigma_X^2 S_X^2/(m S_X^{\star 2}) + \sigma_Y^2 S_Y^2/(n S_Y^{\star 2})}{\sigma_X^2/m + \sigma_Y^2/n}\right)^{1/2},$$

where $(\bar{X}^\star, \bar{Y}^\star, S_X^{\star 2}, S_Y^{\star 2})$ is an independent copy of $(\bar{X}, \bar{Y}, S_X^2, S_Y^2)$. Although $\mathcal{R}(\mathbb{S}, \mathbb{S}^\star, \xi)$ is not itself a FGPQ for θ, $\mathcal{R}_\star(\mathbb{S}, \mathbb{S}^\star, \xi) = (\bar{X} - \bar{Y}) - \mathcal{R}(\mathbb{S}, \mathbb{S}^\star, \xi)$ is a FGPQ for θ. Moreover, it leads to the same confidence interval for θ as does the FGPQ $\mathcal{R}_\theta$ defined by

$$\mathcal{R}_\theta = (\bar{X} - \bar{Y}) - \left[(\bar{X}^\star - \mu_X)\sqrt{\frac{S_X^2}{S_X^{\star 2}}} - (\bar{Y}^\star - \mu_Y)\sqrt{\frac{S_Y^2}{S_Y^{\star 2}}} \right].$$

This latter FGPQ can be derived by a direct application of Theorem 2 of Section 4.

Example 2: Balanced one-way random-effects model. Consider the one-way random-effects model $X_{ij} = \mu + A_i + e_{ij}$, $i = 1, \dots, a$, $j = 1, \dots, n$, where $A_i \sim \mathrm{N}(0, \sigma_A^2)$, $e_{ij} \sim \mathrm{N}(0, \sigma_e^2)$, and $\{A_i\}$ and $\{e_{ij}\}$ are all mutually independent. Define

$$\bar{X}_{i\cdot} = \frac{1}{n}\sum_{j=1}^{n} X_{ij},$$

$$\bar{X}_{\cdot\cdot} = \frac{1}{a}\sum_{i=1}^{a} \bar{X}_i,$$

$$S_B^2 = \frac{n\sum_{i=1}^{a}(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2}{a-1},$$

and

$$S_W^2 = \frac{\sum_{i=1}^{a}\sum_{j=1}^{n}(X_{ij} - \bar{X}_{i\cdot})^2}{a(n-1)}.$$

Thus S_B^2 is the *between-groups* mean square and S_W^2 is the *within-groups* mean square.

Suppose that one is interested in a confidence interval for $\theta = \sigma_A^2$. Approximate confidence interval procedures, such as the Tukey–Williams procedure (Tukey 1951; Williams 1962), have been proposed for this problem. Weerahandi (1993, p. 904) pointed out that a confidence bound for σ_A^2 may be obtained by inverting a generalized test (see Weerahandi 1991, p. 152) of the null hypothesis $H_0 : \sigma_A^2 = \delta$ versus a one-sided alternative. It is easily verified that this process results in a confidence statement of the form $\mathcal{R}_\beta < \sigma_A^2 < \infty$, where $\mathcal{R}_\beta$ is the β-percentile of the conditional distribution of the quantity $\mathcal{R}$ given by

$$\mathcal{R} = \frac{S_B^2}{nS_B^{\star 2}}(\sigma_e^2 + n\sigma_A^2) - \frac{S_W^2}{nS_W^{\star 2}}\sigma_e^2,$$

conditional on S_B^2 and S_W^2, where $(S_B^{\star 2}, S_W^{\star 2})$ is an independent copy of (S_B^2, S_W^2). Observe that $\mathcal{R}$ is a FGPQ for σ_A^2.

Two other FGPQs may be defined that are related to the foregoing FGPQ but are guaranteed to take on only nonnegative values. These are $|\mathcal{R}|$ and $\max(0, \mathcal{R})$. Asymptotically, both of these modified FGPQs are equivalent to $\mathcal{R}$ as long as σ_A^2 and σ_e^2 are nonzero. We expect $\max(0, \mathcal{R})$ to have better small-sample properties.

2.3 Fiducial Generalized Pivotal Quantities and Generalized Test Variables

FGPQs have another useful property. Suppose that θ is a scalar parameter. If $R_\theta(\mathbb{S}, \mathbb{S}^\star, \xi)$ denotes a FGPQ for θ, then the quantity $\theta - R_\theta(\mathbb{S}, \mathbb{S}^\star, \xi)$ is automatically a *generalized test variable* (Tsui and Weerahandi 1989) for testing the null hypothesis $H_0 : \theta \le \theta_0$ versus the alternative $\theta > \theta_0$, where θ_0 is a user-specified value. Results that we prove in this article concerning coverage properties of GCIs can be restated so that they become statements about type I error rates of generalized tests or about the correctness of generalized p values. We do not elaborate further on this here, but a detailed treatment of generalized tests will be the subject of a forthcoming article.

3. MAIN RESULT: FREQUENTIST JUSTIFICATION FOR GENERALIZED CONFIDENCE INTERVALS

As mentioned earlier, the literature abounds with examples of GCIs, but no theoretical results exist that guarantee satisfactory performance of these intervals. The only evidence of their acceptability for practical use is through simulation studies suggesting that almost all reported GCI methods appear to have coverage probabilities close to their stated values.

In the next section we state and prove a theorem that appears to be the first result that guarantees, under some fairly mild conditions, the asymptotic correctness of the coverage probability of a GCI. Although we do not discuss generalized tests here, it is worth noting that, as a corollary to the theorem, one can also establish asymptotic correctness of type I error rates for generalized tests of hypotheses.

Let us consider a parametric statistical problem where we observe $X_1, \dots, X_n$, whose joint distribution belongs to some family of distributions parameterized by $\xi \in \Xi \subset \mathbb{R}^p$. Let $\mathbb{S} = (S_1, \dots, S_k)$ denote a statistic based on the X_i's. In theory, we can consider an independent copy of $X_1^\star, \dots, X_n^\star$ and denote the statistic based on the $X_i^\star$'s by $\mathbb{S}^\star$. Finally, suppose that a function $\mathcal{R}_\theta(\mathbb{S}, \mathbb{S}^\star, \xi)$ is available that is a FGPQ for a scalar parameter $\theta = \pi(\xi)$. We first consider some notation and assumptions.

Assumption A.

1. Assume that there exists $\mathbf{t}(\xi) \in \mathbb{R}^k$ such that

$$\sqrt{n}\big(S_1^\star - t_1(\xi), \dots, S_k^\star - t_k(\xi)\big) \xrightarrow{\mathcal{D}} \mathbf{N} = (N_1, \dots, N_k),$$

where $\mathbf{N}$ has a nondegenerate multivariate normal distribution.

2. Assuming the existence and continuity of second partial derivatives with respect to $\mathbf{s}^\star$ of $\mathcal{R}_\theta(\mathbf{s}, \mathbf{s}^\star, \xi)$, we have the following one-term Taylor expansion with a remainder term:

$$\mathcal{R}_\theta(\mathbf{s}, \mathbb{S}^\star, \xi) = g_{0,n}(\mathbf{s}, \xi) + \sum_{j=1}^{k} g_{1,j,n}(\mathbf{s}, \xi)\big(S_j^\star - t_j(\xi)\big) + R_n(\mathbf{s}, \mathbb{S}^\star, \xi). \quad (1)$$

Here

$$g_{0,n}(\mathbf{s}, \xi) = \mathcal{R}_\theta\big(\mathbf{s}, \mathbf{t}(\xi), \xi\big),$$

$$g_{1,j,n}(\mathbf{s}, \xi) = \frac{\partial}{\partial s_j^\star}\mathcal{R}_\theta(\mathbf{s}, \mathbf{s}^\star, \xi)\Big|_{\mathbf{s}^\star = \mathbf{t}(\xi)},$$

and

$$R_n(\mathbf{s}, \mathbf{s}^\star, \xi) = \sum_{i=1}^{k}\sum_{j=1}^{k}\big(\mathbf{s}_i^\star - t_i(\xi)\big)\big(\mathbf{s}_j^\star - t_j(\xi)\big)\frac{1}{2}\frac{\partial^2}{\partial \mathbf{s}_i^\star \mathbf{s}_j^\star}\mathcal{R}_\theta(\mathbf{s}, \tilde{\mathbf{s}}, \xi),$$

where $\tilde{\mathbf{s}}$ lies on the line segment connecting $\mathbf{t}(\xi)$ and $\mathbf{s}^\star$. Suppose that $\mathcal{A} \subset \mathbb{R}^k$ is an open set containing $\mathbf{t}(\xi)$ with the following properties:

(a) The functions $g_{1,j,n}(\mathbf{s}, \xi)$ converge uniformly in $\mathbf{s} \in \mathcal{A}$ to a function $g_{1,j}(\mathbf{s}, \xi)$ continuous at $\mathbf{s} = \mathbf{t}(\xi)$.

(b) For all $\mathbf{s} \in \mathcal{A}$, $\sum_{j=1}^{k} |g_{1,j}(\mathbf{s}, \xi)| > 0$.

(c) There is $M < \infty$ such that $|\frac{\partial^2}{\partial \mathbf{s}_i^\star \mathbf{s}_j^\star}\mathcal{R}_\theta(\mathbf{s}, \mathbf{s}^\star, \xi)| < M$ for all n and all $(\mathbf{s}, \mathbf{s}^\star) \in \mathcal{A} \times \mathcal{A}$.

We are now ready to state and prove the following theorem, the proof of which is given in Appendix A.

Theorem 1. Suppose that Assumption A holds and that for each fixed $\mathbf{s}$, n, and $\gamma \in (0, 1)$, there exists a real number $C_n(\mathbf{s}, \gamma)$ such that

$$\lim_{n\to\infty} P_\xi\big(\mathcal{R}_\theta(\mathbf{s}, \mathbb{S}^\star, \xi) \le C_n(\mathbf{s}, \gamma)\big) = \gamma. \tag{2}$$

Then $\lim_{n\to\infty} P_\xi(\mathcal{R}_\theta(\mathbb{S}, \mathbb{S}, \xi) \le C_n(\mathbb{S}, \gamma)) = \gamma$. In particular, because $\mathcal{R}_\theta(\mathbb{S}, \mathbb{S}, \xi) = \theta$, it follows that the interval $-\infty < \theta \le C_n(\mathbb{S}, \gamma)$ is a one-sided confidence interval for θ with asymptotic coverage probability equal to γ.

Remark 2. The various conditions stated in Assumption A could be weakened. For example, we do not have to assume that the limiting random variable $\mathbf{N}$ in Assumption A.1 is normal. The proof of Theorem 1 would then have to be modified accordingly. (For an exact statement of the more general version of Assumption A under which Theorem 1 still holds, see Hannig 2005.)

Remark 3. Although we have stated the result of Theorem 1 in terms of coverage probability of a GCI, the same result carries over to generalized p values associated with generalized tests of hypotheses.

Remark 4. The result is easily generalized to the vector parameter case. We state a vector parameter version of Theorem 1 in Appendix B.

Remark 5. The statement of Theorem 1 holds for each fixed ξ. Sometimes it is of interest to establish a stronger conclusion,

$$\lim_{n\to\infty} \sup_{\xi\in\Theta_0} \big|P_\xi\big(\mathcal{R}_\theta(\mathbb{S}, \mathbb{S}, \xi) \le C_n(\mathbb{S}, \gamma)\big) - \gamma\big| = 0, \tag{3}$$

where $\Theta_0 \subset \Theta$. A careful review of the proof of Theorem 1 reveals that (3) holds as long as Assumption A holds uniformly in $\xi \in \Theta_0$.

Example 3: Verification of assumptions of Theorem 1. Consider the two examples of Section 2, the Behrens–Fisher problem and the balanced one-way random model. We may show that the GPQs in these two examples satisfy the assumptions of Theorem 1. Hence the resulting GCIs will have asymptotically correct coverage probabilities. This is, of course, well known for the Behrens–Fisher problem and may be directly verified for the one-way random model. However, our purpose here is to illustrate the application of Theorem 1 by considering some familiar examples. To save space, we illustrate the process for the one-way random model. The conditions may be verified for the Behrens–Fisher problem using analogous arguments.

Proposition 1. In the one-way random model, the $(1 - \alpha)100\%$ GCI for σ_A^2 based on the FGPQ discussed in Section 2 has asymptotically $100(1 - \alpha)\%$ frequentist coverage as $a \to \infty$.

Proof. We need to verify the conditions of Theorem 1. The generalized pivot for σ_A^2 given in Section 2 may be expressed as

$$\mathcal{R}(S_B, S_W, S_B^\star, S_W^\star, \sigma_e^2, \sigma_\alpha^2) = \frac{S_B^2}{nS_B^{\star 2}}(\sigma_e^2 + n\sigma_\alpha^2) - \frac{S_W^2}{nS_W^{\star 2}}\sigma_e^2.$$

We consider n fixed and $a \to \infty$. Combining well-known facts from the theory of linear models with a simple calculation, we get $\sqrt{a}(S_B^{-2} - (\sigma_e^2 + n\sigma_\alpha^2)^{-1}, S_W^{-2} - \sigma_e^{-2}) \xrightarrow{\mathcal{D}} (N_1, N_2)$, where N_1 and N_2 are independent nondegenerate Gaussian random variables. Thus we can set $g_{0,a} = (s_B^2 - s_W^2)/n$, $g_{1,1,a} = s_B^2(\sigma_e^2 + n\sigma_\alpha^2)/n$, $g_{1,2,a} = -s_W^2\sigma_e^2/n$, and $R_n = 0$. The various conditions of Assumption A are immediately verified, and the proposition follows from Theorem 1.

In particular, because $|\mathcal{R}|$ and $\max(0, \mathcal{R})$ are asymptotically equivalent to $\mathcal{R}$ (provided that σ_A^2 and σ_e^2 are nonzero), we may conclude that they also will lead to GCIs with correct asymptotic coverage.

Next we discuss an example where the conditions of the proposition do not hold. It is known that neither fiducial intervals nor GCIs have satisfactory frequentist performance in this example.

Example 4: A situation where conditions of Theorem 1 do not hold. Consider k independent samples $X_{i1}, \dots, X_{in}$ iid $\mathrm{N}(\mu_i, \sigma^2)$, $i = 1, \dots, k$. Suppose that we are interested in one-sided and two-sided confidence intervals for $\theta = \sum_{i=1}^{k} \mu_i^2 = \boldsymbol{\mu}^T\boldsymbol{\mu}$, where $\boldsymbol{\mu} = (\mu_1, \dots, \mu_k)^T$. Let $\mathbb{X}^\star$ be an independent copy of the data vector $\mathbb{X}$. Furthermore, let $\bar{X}_{i\cdot}$ and $\bar{X}_{i\cdot}^\star$ denote the sample means of the ith sample and let S_ν^2 and $S_\nu^{\star 2}$ be the pooled estimates of σ^2 with $\nu = k(n - 1)$ degrees of freedom, based on $\mathbb{X}$ and $\mathbb{X}^\star$, respectively. The obvious generalized pivot (an FGPQ) for θ, given by

$$\mathcal{R}_\theta(\mathbb{X}, \mathbb{X}^\star, \boldsymbol{\mu}, \sigma^2) = \sum_{i=1}^{k}\left(\bar{X}_{i\cdot} + \frac{S_\nu}{S_\nu^\star}(\bar{X}_{i\cdot}^\star - \mu_i)\right)^2, \tag{4}$$

does not lead to intervals for θ with good frequentist properties. This same phenomenon is observed in more general problems involving quadratic functions of fixed effects in mixed linear models. For instance, Daniels et al. (2005), in their work on GCIs for a quadratic function of fixed effects in a two-factor mixed model, were led to consider alternative GPQs because they found that the obvious FGPQ did not lead to satisfactory intervals.

It is easy to see that the assumptions of Theorem 1 are satisfied for the FGPQ in (4) if and only if $\theta > 0$. Toward this end, notice that

$$\mathcal{R}_\theta(\mathbb{X}, \mathbb{X}^\star, \mu, \sigma^2) = \sum_{i=1}^k \bar{X}_{i\cdot}^2 + \sum_{i=1}^k \frac{2S_\nu \bar{X}_{i\cdot}}{\sigma}(\bar{X}_{i\cdot}^\star - \mu_i) + R_n(\mathbb{X}, \mathbb{X}^\star, \boldsymbol{\mu}, \sigma^2),$$

where

$$R_n(\mathbb{X}, \mathbb{X}^\star, \boldsymbol{\mu}, \sigma^2) = \sum_{i=1}^k \left(\frac{S_\nu^2}{S_\nu^{\star 2}}(\bar{X}_{i\cdot}^\star - \mu_i)^2 + 2S_\nu \bar{X}_{i\cdot}(S_\nu^{\star -1} - \sigma^{-1})(\bar{X}_{i\cdot}^\star - \mu_i) \right).$$

Hence all assumptions of Theorem 1 are satisfied as long as $\sigma > 0$ and $\theta > 0$. But if $\theta = 0$, then Assumption A.2(b) is clearly violated. Notice that, regardless of the value of θ, $\mathcal{R}_\theta > 0$ almost surely. Therefore, if $\theta = 0$, then any lower confidence bound based on the distribution of $\mathcal{R}_\theta$ will be positive, and hence will miss the true value $\theta = 0$.

This problem also causes difficulties when one attempts to obtain a fiducial interval for θ. Wilkinson (1977) considered this example in the context of fiducial inference and noted that the point $\boldsymbol{\mu} = \mathbf{0}$ is a special point in the space of $\{\boldsymbol{\mu}\}$. Barnard (1982) also pointed out difficulties that arise when the fiducial distribution of μ is used to derive fiducial distributions for functions of μ that are not one-to-one, such as μ^2. We consider this example further in Example 5, where we propose a new FGPQ that does not suffer the problems at $\theta = 0$ discussed here for the naive FGPQ given in (4).

4. A STRUCTURAL METHOD FOR CONSTRUCTING FIDUCIAL GENERALIZED PIVOTAL QUANTITIES

As mentioned earlier, during the past few years, the idea of GCIs and generalized tests have been used by many authors to obtain useful inference procedures in nonstandard problems. Although Weerahandi (1993) provided a few examples illustrating the application of GPQs, he did not provide a systematic approach for finding GPQs. As a matter of fact, Weerahandi (1993) stated that "the problem of finding an appropriate pivotal quantity is a nontrivial task... [and] ... further research is necessary to develop simple methods of constructing generalized pivotals for classes of general problems, and this is beyond the scope of this article."

Until recently, a general method for constructing GPQs was not available in the literature, and each particular problem appeared to require some ingenuity in constructing an appropriate GPQ or a test variable. Chiang (2001) proposed the method of *surrogate variables* for deriving approximate confidence intervals for functions of variance components in balanced mixed linear models. Iyer and Mathew (2002) pointed out that his intervals are identical to GCIs. Nevertheless, Chiang had indirectly provided a systematic method for computing GCIs for the class of problems that he considered; however, he did not extend his method to any other class of problems, and also did not discuss the connection between his method and either Weerahandi's GCIs or Fisher's fiducial intervals.

In an unpublished technical report, Iyer and Patterson (2002) formulated a method for constructing GPQs and generalized test variables for a parameter $\theta = \pi(\xi)$ and proved that the method works for the class of problems where there exists a k-dimensional pivotal quantity that bears an invertible pivotal relationship with the parameter ξ (see Definition 4). Nearly every GCI in the published statistical literature may be obtained using this recipe, and the construction always yields FGPQs. For instance, Burdick, Borror, and Montgomery (2005) and Burdick, Park, Montgomery, and Borror (2005) reported GCIs for parameters of interest in Gage R&R studies using this recipe. In Theorem 2 we give a reformulation of their construction using the notation of this article.

Because the method underlying Theorem 2, as well as its generalizations given in Theorems 3 and 4, is inspired by, and very much related to, Fraser's (1961, 1966, 1968) development of structural inference, we refer to this method as the *structural method* for constructing FGPQs.

It is useful to first record the following definitions.

Definition 3. Let $\mathbb{S} = (S_1, \dots, S_k) \in \mathcal{S} \subset \mathbb{R}^k$ be a k-dimensional statistic whose distribution depends on a p-dimensional parameter $\xi \in \Xi$. Suppose that there exist mappings $f_1, \dots, f_q$, with $f_j : \mathbb{R}^k \times \mathbb{R}^p \to \mathbb{R}$, such that if $E_i = f_i(\mathbb{S}; \xi)$, for $i = 1, \dots, q$, then $\mathbb{E} = (E_1, \dots, E_q)$ has a joint distribution that is free of ξ. We say that $\mathbf{f}(\mathbb{S}, \xi)$ is a *pivotal quantity* for ξ, where $\mathbf{f} = (f_1, \dots, f_q)$.

Definition 4. Let $\mathbf{f}(\mathbb{S}, \xi)$ be a pivotal quantity for ξ as described in Definition 3 with $q = p$. For each $\mathbf{s} \in \mathcal{S}$, define $\mathcal{E}(\mathbf{s}) = \mathbf{f}(\mathbf{s}, \Xi)$. Suppose that the mapping $\mathbf{f}(\mathbf{s}, \cdot) : \Xi \to \mathcal{E}(\mathbf{s})$ is invertible for every $\mathbf{s} \in \mathcal{S}$. We then say that $\mathbf{f}(\mathbb{S}, \xi)$ is an *invertible pivotal quantity* for ξ. In this case we write $\mathbf{g}(\mathbf{s}, \cdot) = (g_1(\mathbf{s}, \cdot), \dots, g_p(\mathbf{s}, \cdot))$ for the inverse mapping, so that whenever $\mathbf{e} = \mathbf{f}(\mathbf{s}, \xi)$, we have $\mathbf{g}(\mathbf{s}, \mathbf{e}) = \xi$.

The following theorem gives a recipe for constructing FGPQs based on the structural method when an invertible pivotal quantity exists.

Theorem 2. Let $\mathbb{S} = (S_1, \dots, S_k) \in \mathcal{S} \subset \mathbb{R}^k$ be a k-dimensional statistic whose distribution depends on a k-dimensional parameter $\xi \in \Xi$. Suppose that there exist mappings $f_1, \dots, f_k$, with $f_j : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$, such that $\mathbf{f} = (f_1, \dots, f_k)$ is an invertible pivotal quantity with inverse mapping $\mathbf{g}(\mathbf{s}, \cdot)$. Define

$$\begin{aligned}\mathcal{R}_\theta &= \mathcal{R}_\theta(\mathbb{S}, \mathbb{S}^\star, \xi) \\ &= \pi\big(g_1\big(\mathbb{S}, \mathbf{f}(\mathbb{S}^\star, \xi)\big), \dots, g_k\big(\mathbb{S}, \mathbf{f}(\mathbb{S}^\star, \xi)\big)\big) \\ &= \pi\big(g_1(\mathbb{S}, \mathbb{E}^\star), \dots, g_k(\mathbb{S}, \mathbb{E}^\star)\big),\end{aligned}$$

where $\mathbb{E}^\star = \mathbf{f}(\mathbb{S}^\star, \xi)$ is an independent copy of $\mathbb{E}$. Then $\mathcal{R}_\theta$ is a FGPQ for $\theta = \pi(\xi)$. When θ is a scalar parameter, an equal-tailed two-sided $(1 - \alpha)100\%$ GCI for θ is given by $\mathcal{R}_{\theta,\alpha/2} \le \theta \le \mathcal{R}_{\theta,1-\alpha/2}$. Here $\mathcal{R}_{\theta,\gamma} = \mathcal{R}_{\theta,\gamma}(\mathbf{s})$ denotes the 100γth percentile of the distribution of $\mathcal{R}_\theta$ conditional on $\mathbb{S} = \mathbf{s}$. One-sided generalized confidence bounds are obtained in an obvious manner.

Proof. Note that, because the distribution of $\mathbb{E}^\star = \mathbf{f}(\mathbb{S}^\star, \xi)$ does not depend on ξ, the distribution of $\mathcal{R}(\mathbb{S}, \mathbb{S}^\star, \xi)$ given $\mathbb{S} = \mathbf{s}$ does not depend on ξ. In addition, $\mathcal{R}_\theta(\mathbb{S}, \mathbb{S}, \xi) = \theta$ by definition of $\mathbf{g}$ and $\mathbf{f}$. Therefore, $\mathcal{R}_\theta$ satisfies the requirements for it to be a FGPQ for θ.

Remark 6. Weerahandi (2004) described an approach termed the *substitution method*, discussed also by Peterson, Berger, and Weerahandi (2003), for constructing generalized test variables and GPQs in certain classes of problems. This method is essentially the same as the construction given by Iyer and Patterson (2002).

Remark 7. Because $P_{\mathbb{S}}[\theta \le \mathcal{R}_{\theta,\gamma}(\mathbb{S})] = P_{\mathbb{S}}[P_{\mathbb{S}^\star}[\mathcal{R}_\theta(\mathbb{S}, \mathbb{S}^\star, \xi) \le \theta|\mathbb{S}] \le \gamma]$, it follows that a GCI for a parameter θ based on the GPQ $\mathcal{R}_\theta$ is frequentist exact if and only if the distribution of $U(\mathbb{S};\theta) = P_{\mathbb{S}^\star}[\mathcal{R}_\theta(\mathbb{S}, \mathbb{S}^\star, \xi) \le \theta|\mathbb{S}]$ is uniform on $[0, 1]$ for all θ. In particular, $U(\mathbb{S};\theta)$ is an ordinary pivotal quantity for θ. It is also easy to see that $U(\mathbb{S};\theta)$ will have a uniform $[0, 1]$ distribution whenever the inequality $\mathcal{R}_\theta \le \theta$ can be put in the form $q(\mathbb{S}, \xi) \le q(\mathbb{S}^\star, \xi)$ for some function $q(\cdot,\cdot)$. This latter sufficient condition is easy to check in many applications. An interesting example of this was reported by Roy and Mathew (2005), who for a given constant t obtained a generalized confidence bound for $\tau = (t - \mu)/\theta$ based on a right-censored sample from a shifted exponential distribution with density $f(x;\theta,\mu) = (1/\theta)\exp(-(x-\mu)/\theta)I_{[\mu,\infty)}(x)$. They observed that the resulting GCI had exact frequentist coverage. One might note that an ordinary pivotal quantity is available in their problem and is given by the cdf of $T = (t - \hat{\mu})/\hat{\theta}$, where $\hat{\mu}$ and $\hat{\theta}$ are the maximum likelihood estimators (MLEs) for μ and θ.

Remark 8. In the case of models admitting a group structure where the pivotal mapping corresponds to group multiplication, the invertibility of the pivotal relationship is guaranteed because of the invertibility of the group operation. As an illustration, consider the complete sufficient statistic $(\bar{X}, S)$ for the normal family $\mathrm{N}(\mu, \sigma^2)$. Consider the mapping $(E_1, E_2) = \mathbf{f}((\bar{X}, S), (\mu, \sigma)) = ((\bar{X} - \mu)/\sigma, S/\sigma)$. First, note that all three sets $\mathcal{E}$, $\mathcal{X}$, and Ξ can be identified with $\mathbb{R} \times \mathbb{R}^+$, where $\mathbb{R}^+$ is the set of all positive real numbers. For (u_1, u_2) and (v_1, v_2) in $\mathcal{G} = \mathbb{R} \times \mathbb{R}^+$, define a binary operation $\circ$ by $(u_1, u_2) \circ (v_1, v_2) = (u_1 + u_2 v_1, u_2 v_2)$. It is easily verified that $\mathcal{G}$ is a group with this binary operation. The identity element is $(0, 1)$, and the inverse of (u_1, u_2) is $(-u_1/u_2, 1/u_2)$. The pivotal relationship may be expressed as $(E_1, E_2) = (-\mu/\sigma, 1/\sigma) \circ (\bar{X}, S)$. The invertibility of this relationship is guaranteed because $(\bar{X}, S) = (-\mu/\sigma, 1/\sigma)^{-1} \circ (E_1, E_2) = (\mu, \sigma) \circ (E_1, E_2) = (\mu + \sigma E_1, \sigma E_2)$. (For further discussion of this and other related issues, see Fraser 1961; Dawid and Stone 1982.)

We now discuss several examples that illustrate how the structural method of Theorem 2 may be applied to obtain FGPQs in some important applications. The next example is a continuation of Example 4 where the obvious FGPQ did not yield satisfactory GCIs.

Example 5: Continuation of Example 4. We continue with the notation of Example 4. Recall that X_{ij}, $j = 1, \dots, n$, $i = 1, \dots, k$, are k independent samples where $X_{ij} \sim \mathrm{N}(\mu_i, \sigma^2)$. The parameter of interest is $\theta = \sum_{i=1}^k \mu_i^2$. This parameter provides a measure of the extent to which the k normal means deviate from the null hypothesis that the means are all 0. This θ appears in the noncentrality parameter of the distribution of the test statistic for testing this null hypothesis. Note that $\sqrt{\theta}$ is the radius of a sphere centered at $\mathbf{0}$ on whose surface the vector $\boldsymbol{\mu} = (\mu_1, \dots, \mu_k)^T$ lies. We consider the problem of obtaining a lower confidence bound for θ.

For $i = 1, \dots, k$, let $\bar{X}_{i\cdot} = \sum_{j=1}^n X_{ij}/n$, the mean for the ith sample. We define $S_H^2 = \sum_{i=1}^k \bar{X}_{i\cdot}^2$ and $S_W^2 = (\sum_{i=1}^k \sum_{j=1}^n (X_{ij} - \bar{X}_{i\cdot})^2)/(k(n-1))$. We first observe that S_W^2 and S_H^2 are independent. In addition, $((n-1)kS_W^2)/\sigma^2 \sim \chi^2_{k(n-1)}$ and $(nS_H^2)/\sigma^2 \sim \chi^2_{k,n\theta/\sigma^2}$ where $\chi^2_{k,\lambda}$ denotes the noncentral chi-squared distribution with noncentrality parameter λ. We use the definition of noncentrality as given by Rao (1973) that is consistent with the definition used by software packages SAS and R. Some authors define the noncentrality parameter to be half the noncentrality defined by Rao.

Let $F_\nu(x;\lambda)$ denote the value of the noncentral chi-squared cumulative distribution function (cdf) with ν degrees of freedom (df) and noncentrality λ evaluated at x. For $x > 0$, if $F_\nu(x;\lambda) = t \in (0, 1)$, then we write $Q_\nu(t;x) = \lambda$, so that Q_ν is the inverse of F_ν when F_ν is viewed as a function of λ, keeping ν and x fixed. The existence of Q_ν is guaranteed by the monotonicity of F_ν viewed as a function of λ. In what follows it is important to notice that if $t < F_\nu(x; 0)$, then $Q_\nu(t;x) > 0$ is strictly decreasing continuous function, and if $t \ge F_\nu(x; 0)$, then $Q_\nu(t;x) = 0$ by definition. We also write $G_\nu(t;\lambda) = x$ for the inverse of F_ν when F_ν is viewed as a function of x, keeping ν and λ fixed. Define $E_1 = F_k(nS_H^2/\sigma^2; n\theta/\sigma^2)$ and $E_2 = ((n-1)kS_W^2)/\sigma^2$. Clearly, E_1 and E_2 are independent. Furthermore, E_1 has a uniform distribution on $(0, 1)$. Thus it follows that the statistic $\mathbb{S} = (S_W^2, S_H^2)$ and the parameter (θ, σ^2) have an invertible pivotal relationship. Applying the structural method of Theorem 2, we get the following FGPQ for θ:

$$\mathcal{R}_\theta = \frac{\mathcal{R}_{\sigma^2}}{n} Q_k\left(F_k\left(\frac{nS_H^{\star 2}}{\sigma^2}; \frac{n\theta}{\sigma^2}\right); \frac{nS_H^2}{\mathcal{R}_{\sigma^2}}\right),$$

$$\text{where } \mathcal{R}_{\sigma^2} = \sigma^2 \frac{S_W^2}{S_W^{\star 2}}. \qquad (5)$$

As usual, $S_W^{\star 2}$ and $S_H^{\star 2}$ are independent copies of S_W^2 and S_H^2.

It is a known fact (see Johnson and Kotz 1970, chap. 28), that if $\lambda \to \infty$, then the noncentral chi-squared cdf with noncentrality parameter λ approaches the cdf of $\mathrm{N}(k + \lambda, 2(k + 2\lambda))$. The error of the approximation is of the order $O(\lambda^{-1/2})$ uniformly in x. Thus one can verify the conditions of Theorem 1 directly as long as $\theta > 0$.

The more interesting case is $\theta = 0$. Just as before, the conditions of Theorem 1 do not apply. However, this time we can prove directly that a lower confidence bound on θ derived from the FGPQ in (5) has the correct asymptotic coverage. Toward this end, first notice that $U = F_k(nS_H^{\star 2}/\sigma^2; n\theta/\sigma^2)$ has uniform distribution on $(0, 1)$. Then fix α and consider a $100(1-\alpha)\%$ lower confidence bound L. The true value $\theta = 0$ is included in the interval $[L, \infty)$ if and only if $P(\mathcal{R}_\theta(s_H, s_W, S_H^\star, S_W^\star, \xi) = 0) > \alpha$. Notice that $S_W^{\star 2} \xrightarrow{P} \sigma^2$, and therefore $P(\mathcal{R}_\theta(s_H, s_W, S_H^\star, S_W^\star, \xi) = 0) > \alpha$ if and only if $nS_H^2/S_W^2 \le C_n$, where $C_n \to G_k(1-\alpha; 0)$ by Slutsky's theorem. Thus $P_\theta(0 \in [L, \infty)) = P_\theta(nS_H^2/S_W^2 \le C_n) \to P_\theta(nS_H^2/\sigma^2 \le G_k(1-\alpha; 0)) = 1 - \alpha$, establishing that the FGPQ of (5) leads to a generalized lower confidence bound for θ with correct asymptotic coverage.

We have examined the small-sample performance of the foregoing generalized confidence bound in a small simulation study and found the coverage probabilities to be close to the nominal value. The next example has important applications in industrial quality control and quality improvement.

Example 6: Proportion conformance. Consider a population characteristic whose distribution may be assumed to be $\mathrm{N}(\mu, \sigma^2)$ (μ, σ are unknown). Let θ denote the proportion of this population contained in the interval (C_1, C_2), where $C_1 < C_2$ are real numbers (C_1 may be negative infinity and C_2 may be positive infinity). In many practical applications it is of interest to obtain a lower or an upper confidence bound for θ. In the manufacturing of machine parts, a *part* is said to meet specifications provided that its performance characteristic is contained in a prespecified interval (C_1, C_2) determined by engineering requirements. One would like to be assured that θ is large, so a lower confidence bound for θ would be of interest. In environmental applications it is often of interest to ensure that pollutant levels in soil, water, or air are well below an amount determined to be safe by the U.S. Environmental Protection Agency. Here θ may represent the proportion of all possible water samples in which the concentration of arsenic exceeds 5 μg/L. One would like to be assured that this proportion is small and so an upper confidence bound for θ would be of interest. In the case of arsenic concentrations, it may be more reasonable to use a lognormal model, but the problem can be restated using the log scale. The problem of computing confidence bounds for θ has been considered by Chou and Owen (1984) and also by Wang and Lam (1996). These authors have provided methods for computing approximate lower confidence bounds for θ. Here we exhibit a FGPQ for θ derived using the structural method of Theorem 2. Suppose that an iid sample $X_1, \ldots, X_n$ is available from the $\mathrm{N}(\mu, \sigma^2)$ distribution. The parameter of interest is $\theta = \Phi(\frac{C_2-\mu}{\sigma}) - \Phi(\frac{C_1-\mu}{\sigma})$. Let $\mathbb{S} = (\bar{X}, S^2)$, where $\bar{X}$ is the sample mean, $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2/(n-1)$ is the sample variance, and $\xi = (\mu, \sigma^2)$. Applying the structural method, we get $\mathcal{R}_\theta = \Phi(\frac{C_2-\mathcal{R}_\mu}{\mathcal{R}_\sigma}) - \Phi(\frac{C_1-\mathcal{R}_\mu}{\mathcal{R}_\sigma})$ as a FGPQ for θ, where $\mathcal{R}_\sigma = \frac{\sigma S}{S^\star}$ and $\mathcal{R}_\mu = \bar{X} - (\bar{X}^\star - \mu)\frac{S}{S^\star}$. It is straightforward to verify that the assumptions of Theorem 1 are satisfied, and hence the GCIs obtained using the proposed FGPQ have correct asymptotic coverage. Simulation results pertaining to small-sample performance of the GCI for proportion conformance have been reported by Patterson, Hannig, and Iyer (2004b) in an unpublished technical report. They found the performance of the GCIs satisfactory for use in practical applications.

The method of this example is generalizable to more complex situations. Of particular interest is GCIs for the proportion of conformance in the one-way random-effects model. This has also been studied in detail by Patterson et al. (2004b).

The next example illustrates a practical problem arising in pharmaceutical statistics where GCIs provide a useful solution.

Example 7: Average bioequivalence. In bioequivalence studies comparing a test drug to a reference drug, it is of interest to compare the mean responses of the two drugs to ensure that they are (more or less) equally effective. With this in mind, the U.S. Food and Drug Administration (FDA) requires that the lab submitting an approval request demonstrate that certain *equivalence criteria* are satisfied. One such criterion, the *average bioequivalence criterion*, requires that the ratio $\theta = \mu_{\mathrm{T}}/\mu_{\mathrm{R}}$ be sufficiently close to 1, where μ_{T} denotes the mean response to the test drug and μ_{R} denotes the mean for the reference drug. A confidence interval for the ratio $\theta = \mu_{\mathrm{T}}/\mu_{\mathrm{R}}$ is useful in this situation. (Readers interested in details may refer to U.S. FDA 2001.)

A key response variable in such studies is the area under the curve (AUC) relating the plasma drug concentration in a patient to the elapsed time after the drug is administered. In accordance with the FDA guidelines, the analysis of AUC is carried out using the log scale. This is because the distribution of AUC is typically well modeled by a lognormal distribution. So the parameter of interest is the ratio of means of two lognormal distributions. This approach, termed "average bioequivalence," involves the calculation of a 90% confidence interval for the ratio of the averages of test and reference products. To establish bioequivalence, the calculated confidence interval should fall within a bioequivalence limit, usually 80–125% for the ratio of the product averages.

The experimental design of choice in bioequivalence studies is a two-period crossover design with an adequate washout period to minimize carryover effects. But a two-group design (also called a *parallel design*) is the more appropriate design when the half-lives of drugs being tested is very long, and this is recognized in the FDA guidelines. In this example we assume that a two-group design is used for the bioequivalence study.

Let $Y_{ij}, j = 1, \ldots, n_1$, denote independent random variables such that $X_{ij} = \ln(Y_{ij}) \sim \mathrm{N}(\mu_i, \sigma_i^2)$ for $j = 1, \ldots, n_i$, $i = 1, 2$. Then $\theta = \exp(\mu_1 - \mu_2 + \frac{1}{2}(\sigma_1^2 - \sigma_2^2))$ is the ratio of the means of the two lognormal populations. When $\sigma_1 = \sigma_2$, this expression simplifies, and we get $\theta = \exp(\mu_1 - \mu_2)$. The problem of obtaining a confidence interval for θ is straightforward in this special case; however, the problem does not admit an exact solution in the general case.

Let us define, for $i = 1, 2$, $\bar{X}_i = \sum_{j=1}^{n_i} X_{ij}/n_i$ and $S_i^2 = (\sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2)/(n_i - 1)$. The structural method may be applied, and we obtain the following FGPQ for θ:

$$\mathcal{R}_\theta = \exp\biggl(\mathcal{R}_{\mu_1} - \mathcal{R}_{\mu_2} + \frac{1}{2}\bigl(\mathcal{R}_{\sigma_1}^2 - \mathcal{R}_{\sigma_2}^2\bigr)\biggr),$$

where $\mathcal{R}_{\sigma_1} = \sigma_1 S_1/S_1^\star$, $\mathcal{R}_{\sigma_2} = \sigma_2 S_2/S_2^\star$, $\mathcal{R}_{\mu_1} = \bar{X}_1 - (\bar{X}_1^\star - \mu_1)S_1/S_1^\star$, and $\mathcal{R}_{\mu_2} = \bar{X}_2 - (\bar{X}_2^\star - \mu_2)S_2/S_2^\star$. It is straightforward to check that the conditions of Theorem 1 hold. Hence the resulting GCI has correct asymptotic coverage.

We have conducted a simulation study to evaluate the performance of the GCI for θ in small samples. Results of our study indicate that the GCI approach may be recommended for practical applications.

Krishnamoorthy and Mathew (2003) have discussed generalized inference for a single lognormal mean, and Krishnamoorthy (unpublished work) has considered generalized inference for the difference between two lognormal means. The generalized pivotal quantities proposed by them are the same as what one would get by applying the structural method of Theorem 2.

In the next section we generalize the recipe of Theorem 2 in two different ways.

5. TWO GENERALIZATIONS OF THE STRUCTURAL METHOD FOR FIDUCIAL GENERALIZED PIVOTAL QUANTITIES

The recipe given in Theorem 2 is best suited for the case when a complete sufficient statistic is available for the inference problem under consideration, although this is not a prerequisite. First, in the interest of using all of the information in the data, we seek inference methods based on sufficient statistics. It is then appropriate to restrict oneself to procedures based on a minimal sufficient statistic. When this is also complete, the dimension of the minimal sufficient statistic will equal the dimension of the parameter indexing the family of distributions under consideration. We can now directly apply the structural method for constructing FGPQs, provided that we can find an appropriate invertible pivotal relationship between $\mathbb{S}$ and ξ.

When the minimal sufficient statistic is not complete, we may still find a statistic $\mathbb{S}$ that has an invertible pivotal relationship with the parameter $\boldsymbol{\xi}$. In this case one can apply the structural method to get an FGPQ. In this situation there will usually be several choices for the insufficient statistic $\mathbb{S}$ with which to construct an invertible pivotal quantity. Some choices can dominate other choices, but no general guidelines are available to help with such a choice. The following example helps illustrate this situation.

Example 8: Common mean problem. Suppose that $X_{i1}, \ldots, X_{in_i}$ are iid $\mathrm{N}(\mu, \sigma_i^2)$, $i = 1, \ldots, k$. It is of interest to construct a confidence interval for the common mean μ. Let $\bar{X}_i = (\sum_{j=1}^{n_i} X_{ij})/n_i$ and $S_i^2 = (\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2)/(n_i - 1)$. Clearly, $\mathbb{S} = (\bar{X}_1, \ldots, \bar{X}_k, S_1^2, \ldots, S_k^2)$ is a minimal sufficient statistic for $(\mu, \sigma_1^2, \ldots, \sigma_k^2)$, but it is not complete.

Many authors have addressed the problem of point estimation as well as confidence interval estimation for μ in this setting using frequentist approaches based on pivotal quantities. (For some recent work on this problem, see Jordan and Krishnamoorthy 1996; Yu, Sun, and Sinha 1999.) Some of the pivotal quantities considered by these authors include

$$PQ_1 = -2\sum_{i=1}^{k} \log_e(P_i(\mu_0)),$$

$$PQ_2 = \sum_{i=1}^{k} \left| \frac{\bar{X}_i - \mu}{S_i/\sqrt{n_i}} \right|, \qquad \text{and} \tag{6}$$

$$PQ_3 = \sum_{i=1}^{k} \left(\frac{\bar{X}_i - \mu}{S_i/\sqrt{n_i}} \right)^2.$$

In PQ_1, $P_i(\mu_0)$ is the p value for testing $H_0: \mu = \mu_0$ versus $H_a: \mu > \mu_0$ using the Student t test with data from the ith group only. The statistic PQ_1 combines evidence from the k groups following Fisher's method for combining p values (see Fisher 1970). The combined evidence against H_0 is quantified by the area to the right of PQ_1 under a chi-squared-density with k df. Values of μ_0 that lead to a combined p value greater than α form a one-sided confidence interval with upper bound equal to ∞.

Clearly, each $P_i(\mu)$ is a pivotal quantity for μ. So any function of these is also a pivotal quantity for μ. Thus exact confidence regions for μ may be constructed using any one of these pivotal quantities. More generally, whenever multiple pivotal quantities, $P_1, \ldots, P_m$, are available for a parameter θ, any function of the P_i is again a pivotal quantity for θ. The same observations holds for GPQs but not necessarily for FGPQs. But if $P_1, \ldots, P_m$ are FGPQs for θ, then certainly any linear function $\sum_{i=1}^m L_i P_i$, with $\sum_{i=1}^m L_i = 1$, is a FGPQ for θ. We now apply the recipe of Theorem 2. Define $\bar{X}_L = L_1\bar{X}_1 + \cdots + L_k\bar{X}_k$. Let $\mathbb{S} = (S_1^2, \ldots, S_k^2, \bar{X}_L)$. We have the following pivotal relationships:

$$E_i = \frac{(n_i - 1)S_i^2}{\sigma_i^2} \sim \chi^2_{n_i - 1}, \qquad i = 1, \ldots, k,$$

and

$$E_{k+1} = \frac{\bar{X}_L - \mu}{\sqrt{\sum_{i=1}^k L_i^2 \sigma_i^2 / n_i}} \sim \mathrm{N}(0, 1).$$

The conditions needed to apply the structural method of Theorem 2 are satisfied, and we conclude that

$$\mathcal{R}_\mu = \bar{X}_L - (\bar{X}_L^\star - \mu) \frac{\sqrt{\sum_{i=1}^k \sigma_i^2 L_i^2 S_i^2/(n_i S_i^{\star 2})}}{\sqrt{\sum_{i=1}^k \sigma_i^2 L_i^2 / n_i}}$$

is a FGPQ regardless of the choice of $L_1, \ldots, L_k$ as long as they sum to unity.

We revisit this *common mean problem* in Example 9 and again in Example 11 and introduce two more FGPQs. In Section 5 we give the results of a small simulation study comparing selected approaches from the literature and the FGPQs defined in this article.

In the remainder of this section we present two systematic approaches for constructing FGPQs for situations where the minimal sufficient statistic is not complete. The first approach is a two-stage approach for constructing a FGPQ where the structural method of Theorem 2 is invoked at each stage. The second approach offers a very general construction and has connections with fiducial inference and ancillary statistics.

5.1 Two-Stage Construction of a Fiducial Generalized Pivotal Quality

The first alternative is a two-stage approach, which is outlined in the following theorem.

Theorem 3. Let $\xi = (\xi_1, \xi_2)$ and suppose that the following conditions hold:

(a) Assuming that ξ_2 is known, there is a statistic, $\mathbb{S}_1 = \mathbb{S}_1(\xi_2)$, that bears an invertible pivotal relationship with ξ_1; see Definition 4.

(b) There is a statistic $\mathbb{S}_2$ such that $\mathbb{S}_2$ and ξ_2 have an invertible pivotal relationship.

Let $\mathcal{R}_{\xi_1|\xi_2}$ be a FGPQ for ξ_1 obtained based on the pivotal relationship between $\mathbb{S}_1(\xi_2)$ and ξ_1 using the structural method of Theorem 2. Likewise, let $\mathcal{R}_{\xi_2}$ be a FGPQ for ξ_2 obtained using its pivotal relationship with $\mathbb{S}_2$. Then $\mathcal{R}_{\xi_1|\mathcal{R}_{\xi_2}}$ is a FGPQ for ξ_1.

Proof. This theorem follows directly from Theorem 2.

Example 9: Two-stage construction of a FGPQ for the common mean problem. Continuing with the notation of Example 8, we have that $\mathbb{S} = (\bar{X}_1, \ldots, \bar{X}_k, S_1^2, \ldots, S_k^2)$ is a minimal sufficient statistic for $(\mu, \sigma_1^2, \ldots, \sigma_k^2)$, but it is not complete. An independent copy of $\mathbb{S}$ is denoted by $\mathbb{S}^\star = (\bar{X}_1^\star, \ldots, \bar{X}_k^\star, S_1^{\star 2}, \ldots, S_k^{\star 2})$. A realization of $\mathbb{S}$ is denoted by $\mathbf{s} = (\bar{x}_1, \ldots, \bar{x}_k, s_1^2, \ldots, s_k^2)$. In Theorem 3, we take $\xi_1 = \mu$, $\xi_2 = (\sigma_1^2, \ldots, \sigma_k^2)$, $\mathbb{S}_1(\xi_2) = (\frac{n_1\bar{X}_1}{\sigma_1^2} + \cdots + \frac{n_k\bar{X}_k}{\sigma_k^2})/(\frac{n_1}{\sigma_1^2} + \cdots + \frac{n_k}{\sigma_k^2})$, and $\mathbb{S}_2 = (S_1^2, \ldots, S_k^2)$. This leads to the FGPQ given by

$$\mathcal{R}_\mu(\mathbb{S}, \mathbb{S}^\star, \xi) = \frac{n_1\bar{X}_1/\mathcal{R}_{\sigma_1^2} + \cdots + n_k\bar{X}_k/\mathcal{R}_{\sigma_k^2}}{n_1/\mathcal{R}_{\sigma_1^2} + \cdots + n_k/\mathcal{R}_{\sigma_k^2}} - \left(\frac{n_1\bar{X}_1^\star/\mathcal{R}_{\sigma_1^2} + \cdots + n_k\bar{X}_k^\star/\mathcal{R}_{\sigma_k^2}}{n_1/\mathcal{R}_{\sigma_1^2} + \cdots + n_k/\mathcal{R}_{\sigma_k^2}} - \mu\right).$$

Proposition 2. Let all $n_1, \ldots, n_k$ approach infinity in such a way that $c_j = \lim n_j/(n_1 + \cdots + n_k)$ exists and $0 < c_j < 1$. Then the $100(1-\alpha)\%$ confidence interval for μ based on the FGPQ $\mathcal{R}_\mu$ has asymptotically $100(1-\alpha)\%$ frequentist coverage.

The proof of this proposition is given in Appendix A. Next, we describe a second alternative construction of FGPQs. This is a more general approach that in principle is applicable to any parametric problem, but the properties of the resulting GCIs are much more difficult to assess. Nonetheless, we can show that GCIs obtained using this approach have exact frequentist coverage in some special situations.

5.2 A General Recipe for Constructing Fiducial Generalized Pivotal Quantities

The two-stage construction is a systematic approach for obtaining a FGPQ for a number of problems where the minimal sufficient statistic is not complete. In many situations it is fairly straightforward to verify that the conditions of Theorem 1 hold for the two-stage FGPQ so that the resulting interval will have, at least asymptotically, the correct frequentist coverage. However, the two-stage construction is perhaps not general enough; see Example 11. In this section we provide a general method for constructing FGPQs and illustrate its application through examples.

Let $\mathbb{S} = (S_1, \ldots, S_k) \in \mathcal{S} \subset \mathbb{R}^k$ denote a statistic whose distribution is indexed by $\xi \in \Xi \subset \mathbb{R}^p$. Let $\theta = \pi(\xi)$ be the parameter of interest. Estimability considerations suggest that it is reasonable to restrict ourselves to cases where $k \ge p$. When $k = p$, the structural method of Theorem 2 is applicable. In this section we outline a general construction for obtaining FGPQs in situations where $p < k$ (more statistics than parameters). We call this the *general structural method.* This construction reduces to the structural method of Theorem 2 when $p = k$.

We make the following assumption.

Assumption B. (a) There exists a mapping $\mathbf{f}: \mathcal{S} \times \Xi \to \mathbb{R}^k$ such that $\mathbb{E} = \mathbf{f}(\mathbb{S}, \xi)$ has a continuous cdf that does not depend on ξ. Let $\mathcal{E} = \mathbf{f}(\mathcal{S} \times \Xi)$. Write $\mathbf{f} = (f_1, \ldots, f_k)$, so that $E_i = f_i(\mathbb{S}, \xi)$ for $i = 1, \ldots, k$.

(b) Let $\mathbb{E}_0 = (E_1, \ldots, E_p)$ and $\mathbf{f}_0 = (f_1, \ldots, f_p)$. We assume that for each fixed $\mathbf{s} \in \mathcal{S}$, the mapping $\mathbf{f}_0(\mathbf{s}, \cdot): \Xi \to \mathcal{E}$ defined by $\mathbf{e}_0 = \mathbf{f}_0(\mathbf{s}, \xi)$ is invertible. Write $\mathcal{E}_0 = \mathbf{f}_0(\mathbf{s}, \Xi)$. We denote this inverse mapping from $\mathcal{E}_0$ to Ξ by $\mathbf{g}_0(\mathbf{s}, \cdot) = (g_1(\mathbf{s}, \cdot), \ldots, g_p(\mathbf{s}, \cdot))$; thus we have $\mathbf{g}_0(\mathbf{s}, \mathbf{f}_0(\mathbf{s}, \xi)) = \xi$ for each $\mathbf{s} \in \mathcal{S}$.

Remark 9. It is not necessary that $\mathbb{E}_0$ consist of the first p elements of $\mathbb{E}$, but we can always achieve this by relabeling if necessary. Note that condition (a) simply says that $\mathbf{f}(\mathbb{S}, \xi)$ is a pivotal quantity for ξ with a continuous cdf, and condition (b) is a *partial invertibility* condition similar to what was required in Definition 4 for a pivotal quantity to be invertible.

Now let $\mathbb{E}_c = (E_{p+1}, \ldots, E_k)$ and $\mathbf{f}_c = (f_{p+1}, \ldots, f_k)$. Substituting $\xi = \mathbf{g}_0(\mathbb{S}, \mathbb{E}_0)$ in the equations $E_j = f_j(\mathbb{S}, \xi)$, $j = p+1, \ldots, k$, we get the identity

$$\mathbb{E}_c = \mathbf{f}_c\big(\mathbb{S}, \mathbf{g}_0(\mathbb{S}, \mathbb{E}_0)\big). \tag{7}$$

For any fixed $\mathbf{s} \in \mathcal{S}$, let $\mathcal{M}(\mathbf{s})$ denote the set of $\mathbf{e} = (e_1, \ldots, e_k) \in \mathcal{E}$ satisfying

$$e_j = f_j\big(\mathbf{s}, \mathbf{g}_0(\mathbf{s}, \mathbf{e}_0)\big), \qquad j = p+1, \ldots, k, \tag{8}$$

where $\mathbf{e}_0 = (e_1, \ldots, e_p)$, the first p coordinates of $\mathbf{e}$. Thus $\mathcal{M}(\mathbf{s})$ is a manifold in $\mathbb{R}^k$.

Given $\mathbb{S}$, by virtue of (7), it follows that $\mathbb{E}$ must lie on the manifold $\mathcal{M}(\mathbb{S})$. Note that the same manifold may be definable using a different set of equations. We use the notation $\mathbb{B}(\mathbf{s}, \mathbf{e}) = \mathbf{0}$ for the equations chosen to define $\mathcal{M}(\mathbf{s})$. In some situations it may be possible to express the equations defining $\mathcal{M}(\mathbf{s})$ in the form $\mathbf{a}(\mathbf{s}) - \mathbf{b}(\mathbf{e}) = \mathbf{0}$ for suitably chosen mappings $\mathbf{a}$ and $\mathbf{b}$. Clearly, $\mathbf{a}(\mathbb{S})$ is an ancillary statistic, because its distribution is not dependent on ξ.

We now make the following assumption concerning the random vector $\mathbb{E}_0$.

Assumption C. Conditional on $\mathbb{B}(\mathbf{s}, \mathbb{E})$, $\mathbb{E}_0$ has a jointly continuous distribution.

The following lemma states the multivariate version of the probability integral transform.

Lemma 1. Let $F_1(\cdot|\mathbb{B}(\mathbf{s}, \mathbb{E}))$ denote the cdf of E_1 given $\mathbb{B}(\mathbf{s}, \mathbb{E})$ and $F_j(\cdot|E_1, \ldots, E_{j-1}, \mathbb{B}(\mathbf{s}, \mathbb{E}))$ denote the cdf of E_j given $E_1, \ldots, E_{j-1}, \mathbb{B}(\mathbf{s}, \mathbb{E})$, $j = 2, \ldots, p$. The conditional joint distribution of $\mathbb{E}_0$ given $\mathbb{B}(\mathbf{s}, \mathbb{E})$ is completely determined by the univariate cdf's $F_1, \ldots, F_p$. Let the random vector $\mathbb{U}$ be defined by

$$\begin{aligned}\mathbb{U} &= \mathbf{F}\big(\mathbb{E}_0|\mathbb{B}(\mathbf{s}, \mathbb{E})\big)\\ &= \Big(F_1\big(E_1|\mathbb{B}(\mathbf{s}, \mathbb{E})\big), F_2\big(E_2|E_1, \mathbb{B}(\mathbf{s}, \mathbb{E})\big), \ldots,\\ &\qquad F_p\big(E_p|E_1, \ldots, E_{p-1}, \mathbb{B}(\mathbf{s}, \mathbb{E})\big)\Big).\end{aligned} \tag{9}$$

The distribution of $\mathbb{U}$, conditional on $\mathbb{B}(\mathbf{s}, \mathbb{E})$, is uniform on $[0, 1]^p$. Hence this is also its unconditional distribution.

We denote the inverse of $\mathbf{F}(\cdot|\mathbb{B}(\mathbf{s}, \mathbb{E}))$ by $\mathbf{G}(\cdot|\mathbb{B}(\mathbf{s}, \mathbb{E}))$. Both $\mathbf{F}$ and $\mathbf{G}$ depend on ξ, but for clarity of notation we do not explicitly display this dependence.

We now state our general construction for FGPQs.

Theorem 4. Suppose that Assumptions B and C are satisfied. Let $\theta = \pi(\xi)$ be a scalar parameter. Define

$$\mathcal{R}_\theta = \mathcal{R}_\theta(\mathbb{S}, \mathbb{S}^\star, \xi) = \pi\big(\mathbf{g}_0\big(\mathbb{S}, \mathbf{G}\big(\mathbf{F}\big(\mathbf{f}_0(\mathbb{S}^\star, \xi)|\mathbb{B}(\mathbb{S}, \mathbb{E}^\star)\big)\big|\mathbf{0}\big)\big)\big).$$

Then $\mathcal{R}_\theta$ is an FGPQ for θ.

Proof. This is obvious by construction.

Remark 10. By the definition of $\mathbf{F}$ and $\mathbf{G}$, it follows that the distribution of

$$\mathbf{G}\big(\mathbf{F}\big(\mathbf{f}_0(\mathbb{S}^\star, \xi)|\mathbb{B}(\mathbb{S}, \mathbb{E}^\star)\big)\big|\mathbf{0}\big)$$

is the same as the distribution of $\mathbb{E}_0^\star$ given $\mathbb{B}(\mathbb{S}, \mathbb{E}^\star) = 0$. We use this fact in the proof of Theorem 5.

Note that the FGPQ defined here will generally depend on the choice of the defining equations for $\mathcal{M}(\mathbf{s})$. An example of this phenomenon is given later in this section. An obvious question now is whether a theorem analogous to Theorem 1 may be established that would guarantee correct asymptotic coverage for GCIs based on this general construction of FGPQs. We conjecture that such a result exists under fairly general conditions, but we do not have an actual theorem at this time. This conjecture is supported by the examples that we have considered. In principle, we should be able to use Theorem 1 directly as long as we can get the conditional cdf in closed form. However, there are problems caused by the fact that the statistics $\mathbb{B}$ often converge to the edge of their sample space. This causes various quantities in Assumptions A to be unbounded. We might be able to overcome this difficulty by paying special attention to the quantities $\mathbb{B}(\mathbb{S}, \mathbb{E}^\star)$ used in the determination of conditional distributions.

Example 10, discussed shortly, provides an illustration of the difficulties outlined in the previous paragraph. This example also illustrates a situation where one would need to use a normalization factor n rather than $\sqrt{n}$; in this connection, see the technical report of Hannig (2005).

5.3 Exact Coverage of Generalized Confidence Intervals in Special Cases

For the case where $k = p$, Remark 7 pointed out that GCIs can have exact frequentist coverage under certain conditions. When $p = 1$ (i.e., when ξ is a scalar parameter), one can show that a generalized confidence region for ξ has exact frequentist coverage under certain circumstances even when $k > 1$. This is the content of the next theorem, whose proof is based on ideas given by Dawid and Stone (1982).

Theorem 5. Suppose that ξ is a scalar parameter. Assume that the manifold $\mathcal{M}(\mathbb{S})$ can be defined using an equation of the form $\mathbf{a}(\mathbf{s}) = \mathbf{b}(\mathbf{e})$, and let $R(\mathbb{S}, \mathbb{S}^\star, \xi)$ denote the FGPQ obtained using the general construction. Let $P[\cdot]$ denote the probability measure associated with $\mathbb{S}$ and let $P_\star[\cdot]$ denote the measure associated with $\mathbb{S}^\star$. Let $\mathcal{C}(\mathbb{S}, \beta)$ be a set in $\boldsymbol{\Xi}$ satisfying $P_\star[\mathcal{R}(\mathbf{s}, \mathbb{S}^\star, \xi) \in \mathcal{C}(\mathbf{s}, \beta)] = \beta$. Let $\mathcal{A}(\mathbf{s}, \beta)$ denote the image of $\mathcal{C}(\mathbf{s}, \beta)$ under the mapping $\mathbf{f}_0(\mathbf{s}, \cdot)$. Assume that $\mathcal{A}(\mathbf{s}_1, \beta) = \mathcal{A}(\mathbf{s}_2, \beta)$ whenever $\mathbf{a}(\mathbf{s}_1) = \mathbf{a}(\mathbf{s}_2)$. Then

$$P[\xi \in \mathcal{C}(\mathbb{S}, \beta)] = \beta.$$

Namely, the generalized confidence region $\mathcal{C}(\mathbb{S}, \beta)$ for ξ with confidence coefficient β has frequentist coverage also equal to β.

Proof. We have

$$\begin{aligned}
\beta &= P_\star[\mathcal{R}(\mathbf{s}, \mathbb{S}^\star, \xi) \in \mathcal{C}(\mathbf{s}, \beta)],\\
&\quad \text{by the definition of } \mathcal{C}(\mathbf{s}, \beta);\\
&= P_\star\big[\mathbf{g}_0(\mathbf{s}, \mathbb{E}_0^\star) \in \mathcal{C}(\mathbf{s}, \beta)|\mathbb{B}(\mathbf{s}, \mathbb{E}^\star) = 0\big],\\
&\quad \text{by the definition of } \mathbf{F} \text{ and } \mathbf{G} \text{ (see Remark 10)};\\
&= P_\star\big[\mathbb{E}_0^\star \in \mathcal{A}(\mathbf{s}, \beta)|a(\mathbf{s}) = b(\mathbb{E}^\star)\big],\\
&\quad \text{by assumption about the form of } \mathbb{B}(\mathbf{s}, \mathbf{e});\\
&= P\big[\mathbb{E}_0 \in \mathcal{A}(\mathbf{s}, \beta)|a(\mathbf{s}) = b(\mathbb{E})\big],\\
&\quad \text{because } \mathbb{E} \text{ is an independent copy of } \mathbb{E}^\star;\\
&= P\big[\mathbb{E}_0 \in \mathcal{A}(\mathbf{s}, \beta)|a(\mathbf{s}) = a(\mathbb{S})\big],\\
&\quad \text{because } b(\mathbb{E}) = a(\mathbb{S}) \text{ is an identity};\\
&= P\big[\mathbf{f}_0(\mathbb{S}, \xi) \in \mathcal{A}(\mathbb{S}, \beta)|a(\mathbf{s}) = a(\mathbb{S})\big],\\
&\quad \text{by definition of } \mathbb{E}_0 \text{ and assumption about } \mathcal{A}(\mathbb{S}, \beta);\\
&= P\big[\xi \in \mathcal{C}(\mathbb{S}, \beta)|a(\mathbf{s}) = a(\mathbb{S})\big],\\
&\quad \text{because } \mathbf{g}_0(\mathbb{S}, \mathbb{E}_0) = \xi \text{ and } \mathbf{g}_0(\mathbb{S}, \mathcal{A}(\mathbb{S}, \beta)) = \mathcal{C}(\mathbb{S}, \beta).
\end{aligned}$$

Therefore, it follows that $P[\xi \in \mathcal{C}(\mathbb{S}, \beta)] = \beta$.

We now illustrate the general construction of this section using two examples. We also use the first example to demonstrate that different choices of equations for defining $\mathcal{M}(\mathbb{S})$ may result in different (nonequivalent) FGPQs. But in this example they turn out to be asymptotically equivalent, thus leading us to conjecture that under certain conditions, different choices of $\mathbb{B}(\mathbf{s}, \mathbb{E})$ may lead to asymptotically equivalent GCIs. The second example is an application of the foregoing general construction to the common mean problem, which yields an FGPQ different from that obtained from the two-stage construction discussed earlier.

Example 10: Nonuniqueness of FGPQ arising due to different choices of $\mathbb{B}(\mathbb{S}, \mathbb{E})$. Suppose that $\mathbb{E} = (E_M, E_m)$ and $\mathbb{S} = (X_M, X_m)$. We assume that $X_i = \theta E_i$, $i = M, m$, $0 < \theta \in \mathbb{R}$, E_M and E_m are distributed as the maximum and the minimum of n iid uniform$[1, 2]$ random variates. We wish to obtain a FGPQ for θ. We now show that two different FGPQs arise from two different choices of defining equations for $\mathcal{M}(\mathbb{S})$, but the resulting GCIs both have asymptotically correct coverage.

Take $\mathbf{f}(\mathbb{S}, \theta) = (E_M, E_m) = \mathbb{E}$, where $E_M = X_M/\theta$ and $E_m = X_m/\theta$. Clearly, the distribution of $\mathbb{E}$ is free of θ. Let $\mathbb{E}_0 = E_M$, so that $\mathbf{f}_0(\mathbb{S}, \xi) = X_M/\theta$. The mapping $\mathbf{g}_0$ may be defined by $\theta = \mathbf{g}_0(\mathbb{S}, \mathbb{E}_0) = X_M/E_M$. Substituting this in the relation $E_m = X_m/\theta$ gives the identity $E_M/E_m = X_M/X_m$. Let $\mathbf{s} = (x_M, x_m)$ and $\mathbf{e} = (e_M, e_m)$ denote realizations of $\mathbb{S}$ and $\mathbb{E}$. Write $a(\mathbf{s}) = x_M/x_m$ and $u(\mathbf{e}) = e_M/e_m$. The manifold (curve) $\mathcal{M}(\mathbf{s})$ on which $\mathbf{e}$ must lie may be described by the equation $\mathbb{B}(\mathbf{s}, \mathbf{e}) = a(\mathbf{s}) - u(\mathbf{e}) = 0$. The same manifold may also be described using the condition $\mathbb{B}(\mathbf{s}, \mathbf{e}) = e_M - ae_m = 0$, where $a = x_M/x_m$. It is easy to check that one gets nonequivalent FGPQs depending on the choice of $\mathbb{B}$ for defining $\mathcal{M}(\mathbf{s})$. We state only the results here omitting the detailed calculations.

- *FGPQ using* $\mathbb{B}(\mathbf{s},\mathbf{e}) = e_M/e_m - a$. Applying the general construction, we get the following generalized upper confidence bound for θ:

$$\theta \le \left[\frac{X_M^n}{\alpha 2^n + (1-\alpha)A^n}\right]^{1/n},$$

where $A = X_M/X_m$. This is the same upper confidence bound obtained by considering the conditional distribution of the pivot X_M/θ conditional on the ancillary statistic $A = X_M/X_m$. Thus by Theorem 5, this interval has correct frequentist coverage $1-\alpha$, conditionally on A, and hence also unconditionally.

- *FGPQ using* $\mathbb{B}(\mathbf{s},\mathbf{e}) = e_M - ae_m$. Here the general construction leads to the upper confidence bound given by

$$\theta < \left[\frac{X_M^{n-1}}{\alpha 2^{n-1} + (1-\alpha)A^{n-1}}\right]^{1/(n-1)},$$

where $A = X_M/X_m$. Unlike the earlier fiducial interval, which was based on conditioning on an ancillary statistic, this fiducial interval does not have exact frequentist coverage. Although it is based on the pivotal quantity

$$W = \frac{(X_M/\theta)^{n-1}}{\alpha 2^{n-1} + (1-\alpha)A^{n-1}},$$

it uses an incorrect percentile of the distribution of the pivot. This is seen by noting that the fiducial interval is obtained from the probability statement $\Pr[W > 1] = 1-\alpha$, but the value 1 used as the α-percentile of W is incorrect. Instead, using the exact probability statement

$$\Pr\left[W > \frac{[\alpha + (A/2)^n(1-\alpha)]^{(n-1)/n}}{[\alpha + (A/2)^{n-1}(1-\alpha)]}\right]$$

would lead to an exact frequentist confidence interval. This observation also helps us demonstrate that the fiducial interval has the correct frequentist coverage asymptotically, because as $n \to \infty$

$$\frac{[\alpha + (A/2)^n(1-\alpha)]^{(n-1)/n}}{[\alpha + (A/2)^{n-1}(1-\alpha)]} \to 1,$$

and W converges in distribution to a nondegenerate limit.

The issue of nonuniqueness observed in this example is related to the Borel paradox described by, for example, Casella and Berger (2002, sec. 4.9.3). (For a more in-depth discussion of the various types of nonuniqueness associated with conditional distributions, see Hannig 1996.)

Example 11: FGPQ for the common mean problem using the general construction. For a less trivial illustration of the general construction for FGPQ, we once again consider the common mean problem, the subject of Examples 8 and 9. The FGPQ for μ in Example 9, obtained using the two-stage construction method, is by no means unique. An application of the general construction yields a different FGPQ for μ.

Consider the invertible pivotal relationship given by $\bar{X}_i = \mu + \sqrt{\sigma_i^2/n_i}E_i$, $(n_i-1)S_i^2 = \sigma_i^2 E_{k+i}$, for $i = 1,\dots,k$, where for $i = 1,\dots,k$, $E_i \sim \mathrm{N}(0,1)$, $E_{k+i} \sim \chi^2_{n_i-1}$, and E_i, $i = 1,\dots,2k$ are jointly independent. Let $\mathbb{S} = (\bar{X}_1,\dots,\bar{X}_k, S_1^2,\dots,S_k^2)$ and $\mathbb{E} = (E_1,\dots,E_{2k})$. Let $\mathbf{s} = (\bar{x}_1,\dots,\bar{x}_k, s_1^2,\dots,s_k^2)$ denote a realization of $\mathbb{S}$, and let $\mathbf{e}$ be the corresponding realization of $\mathbb{E}$. Then $\mathbf{e}$ must take values on the manifold $\mathcal{M}(\mathbf{s})$ defined by the $k-1$ equations

$$\begin{aligned} 0 &= \frac{e_i}{\sqrt{e_{k+i}/(n_i-1)}}\sqrt{\frac{s_i^2}{n_i}} - \frac{e_k}{\sqrt{e_{2k}/(n_k-1)}}\sqrt{\frac{s_k^2}{n_k}} - (\bar{x}_i - \bar{x}_k) \\ &= t_i\sqrt{\frac{s_i^2}{n_i}} - t_k\sqrt{\frac{s_k^2}{n_k}} - (\bar{x}_i - \bar{x}_k) \quad \text{for } i = 1,\dots,k-1, \end{aligned}$$

where t_i, $i = 1,\dots,k$, are realizations of independent Student t random variables with $n_i - 1$ degrees of freedom. For $i = 1,\dots,k-1$, define $B_i = T_i\sqrt{s_i^2/n_i} - T_k\sqrt{s_k^2/n_k} - (\bar{x}_i - \bar{x}_k)$. A fiducial distribution for μ may be defined as the conditional distribution of $\mu = \bar{x}_k - T_k\sqrt{s_k^2/n_k}$ given $B_1 = B_2 = \cdots = B_{k-1} = 0$ (see, e.g., Fisher 1961a,b). Let $\mathbb{B} = (B_1,\dots,B_{k-1})$ and $\mathbf{b} = (b_1,\dots,b_{k-1})$. First, we note that the joint probability density function (pdf) of $B_1,\dots,B_{k-1},\mu$ is given by

$$\begin{aligned} &f_{B_1,\dots,B_{k-1},\mu}(e_1,\dots,e_{k-1},e_k) \\ &\quad = K\left(\left(1 + \frac{(\bar{x}_1 - e_k - e_1)^2}{\hat{\sigma}_1^2}\right)^{n_1/2} \cdots \right. \\ &\qquad \times \left(1 + \frac{(\bar{x}_{k-1} - e_k - e_{k-1})^2}{\hat{\sigma}_{k-1}^2}\right)^{n_{k-1}/2} \\ &\qquad \left. \times \left(1 + \frac{(\bar{x}_k - e_k)^2}{\hat{\sigma}_k^2}\right)^{n_k/2}\right)^{-1}, \end{aligned}$$

where K is the normalizing constant and $\hat{\sigma}_i^2 = (n_i-1)s_i^2/n_i$ is the MLE of σ_i^2. Hence the conditional pdf of μ given $B_i = b_i$ for $i = 1,\dots,k-1$, is

$$\begin{aligned} f_{\mu|b_1,\dots,b_{k-1}}(t) &= C\left[\left(1 + \frac{(\bar{x}_1 - t - b_1)^2}{\hat{\sigma}_1^2}\right)^{n_1/2} \cdots \right. \\ &\quad \times \left(1 + \frac{(\bar{x}_{k-1} - t - b_{k-1})^2}{\hat{\sigma}_{k-1}^2}\right)^{n_{k-1}/2} \\ &\quad \left. \times \left(1 + \frac{(\bar{x}_k - t)^2}{\hat{\sigma}_k^2}\right)^{n_k/2}\right]^{-1}, \end{aligned}$$

where

$$\begin{aligned} C^{-1} &= \int_{-\infty}^{\infty} dt \left[\left(1 + \frac{(\bar{x}_1 - t - b_1)^2}{\hat{\sigma}_1^2}\right)^{n_1/2} \cdots \right. \\ &\quad \times \left(1 + \frac{(\bar{x}_{k-1} - t - b_{k-1})^2}{\hat{\sigma}_{k-1}^2}\right)^{n_{k-1}/2} \\ &\quad \left. \times \left(1 + \frac{(\bar{x}_k - t)^2}{\hat{\sigma}_k^2}\right)^{n_k/2}\right]^{-1}. \end{aligned}$$

Setting $b_i = 0$ for $i = 1,\dots,k-1$, yields the required fiducial pdf for μ. This is given by

$$\begin{aligned} f(t) = C\Bigg[&\left(1 + \frac{(\bar{x}_1 - t)^2}{\hat{\sigma}_1^2}\right)^{n_1/2} \cdots \left(1 + \frac{(\bar{x}_{k-1} - t)^2}{\hat{\sigma}_{k-1}^2}\right)^{n_{k-1}/2} \\ &\times \left(1 + \frac{(\bar{x}_k - t)^2}{\hat{\sigma}_k^2}\right)^{n_k/2}\Bigg]^{-1}. \end{aligned} \tag{10}$$

A FGPQ $\mathcal{R}_\mu$ corresponding to this fiducial density may be obtained using the general construction. However, it is quite convenient to use the fiducial framework when deriving the actual GCI. The FGPQ in this example does not appear to satisfy the conditions of Theorem 1, but we can still prove directly that confidence intervals based on $\mathcal{R}_\mu$ are asymptotically correct. This is the message of the next proposition, the proof of which is given in Appendix A.

Proposition 3. Let all $n_1, \dots, n_k$ approach infinity in such a way that $c_j = \lim n_j/(n_1 + \cdots + n_k)$ exists and $0 < c_j < 1$. Then the $100(1-\alpha)\%$ confidence interval for μ based on the FGPQ $\mathcal{R}_\mu$ corresponding to the fiducial density in (10) has asymptotically $100(1-\alpha)\%$ frequentist coverage.

The unpublished technical report of Patterson et al. (2004a) gives the results of a detailed simulation study comparing three different procedures for constructing confidence intervals for the common mean problem: Fisher's method [PQ_1 in (6)], GCI using the two-stage FGPQ, and the GCI using the general FGPQ construction. Fisher's method was chosen based on the results of Yu et al. (1999), which indicate that Fisher's method performs as well as or better than competing procedures in most situations.

Table 1 gives a representative subset of the results from Patterson et al. (2004a). We report the case of $k = 3$ (three laboratories) for selected sample sizes and standard deviations. The first row for each scenario gives the empirical coverage probability for the upper confidence bound corresponding to a stated coverage of .95. The second row gives the empirical expected value of the absolute distance of the upper bound from the true parameter value for each of the three methods. Smaller values of the expected absolute distance imply greater efficiency of the method. In Table 1 this expected absolute distance is referred to as "expected length."

Fisher's method is an exact method, but the two GCI methods are only asymptotically exact. Nevertheless, we see that the small-sample performance of the two GCI methods are quite satisfactory, and they appear to be more efficient than Fisher's method. We also see that the GCI using the general construction performs slightly better than the two-stage GCI in terms of confidence interval expected length.

6. CONNECTION BETWEEN A FIDUCIAL GENERALIZED PIVOTAL QUANTITY FOR A PARAMETER AND ITS FIDUCIAL DISTRIBUTION

Fraser (1961) considered fiducial inference for a normal mean μ with unit variance and outlined a method for providing a frequency interpretation for the fiducial distribution of μ. He stated the following:

> Let $\mu^\star$ designate possible values for the parameter relative to an observed $\bar{x}$;.... The statistical problem admits free translation on the $\bar{x}$ axis. Consider a very large number of samples from normal distributions with the specifications of this example. In each case translate the sample mean.... The parameter values will be correspondingly translated. Simple mathematics then shows that the frequency distribution of these translated means $\mu^\star$ is normal with center at $\bar{x}$ and with unit scale parameter. There is thus a frequency distribution of parameter values $\mu^\star$ that might have produced the observed $\bar{x}$.

What Fraser proposed in his 1961 article could be written, in the notation of GPQs, as

$$\mathcal{R}_\mu = \mathcal{R}_\mu(\mathbb{S}, \mathbb{S}^\star, \xi) = \bar{X} - (\bar{X}^\star - \mu). \qquad (11)$$

This becomes clear by noting that when Fraser considers "a very large number of samples with the specifications of this example," he is actually considering the random variable $\bar{X}^\star$, which is an independent copy of $\bar{X}$. His *translated means* are produced by translating $\bar{X}$ by an amount equal to $E^\star = \bar{X}^\star - \mu$. Viewed thusly, Fraser's $\mu^\star$ is exactly the $\mathcal{R}_\mu$ defined in (11). When he describes the distribution of $\mu^\star$, he is essentially describing the distribution of $\mathcal{R}_\mu$ and the frequency interpretation for the fiducial distribution of a parameter is then automatic.

Table 1. Empirical Coverages and Expected Lengths for Fisher's Exact Upper Confidence Bound and Two Generalized Upper Bounds

n_1	n_2	n_3	σ_1	σ_2	σ_3	Two-stage	General FGPQ	Fisher
7	20	12	1.00	1.00	1.00	.9375	.9390	.9510
						(.2791)	(.2769)	(.3025)
7	20	12	1.00	.10	16.00	.9450	.9440	.9510
						(.1231)	(.1216)	(.1546)
10	10	10	1.00	1.00	1.00	.9345	.9340	.9480
						(.3203)	(.3189)	(.3460)
50	50	50	1.00	1.00	1.00	.9470	.9447	.9490
						(.1398)	(.1393)	(.1473)
50	50	50	1.00	.10	10.00	.9460	.9455	.9490
						(.0728)	(.0724)	(.0866)
40	30	50	1.00	.10	10.00	.9485	.9483	.9493
						(.0930)	(.0925)	(.1095)
40	30	50	1.00	10.00	.10	.9490	.9485	.9517
						(.0731)	(.0726)	(.0957)
7	20	8	1.00	.25	1.00	.9435	.9443	.9485
						(.1812)	(.1785)	(.2048)
100	80	90	1.00	.25	.25	.9517	.9503	.9527
						(.0602)	(.0598)	(.0643)
100	80	90	1.00	16.00	.20	.9517	.9505	.9527
						(.0717)	(.0712)	(.0849)
100	100	100	1.00	1.00	1.00	.9530	.9527	.9525
						(.0974)	(.0969)	(.1020)

NOTE: In each of the last three columns, two numbers are reported for each combination of sample sizes and standard deviations. The first of these numbers is the empirical coverage probability, corresponding to a claimed coverage of .95. The second number, given in parentheses, is the empirical average absolute distance from the upper confidence bound to the true mean.

In general, FGPQs allow us to associate a distribution with a parameter θ akin to a fiducial distribution for θ. For example, when considering an upper confidence bound for a scalar parameter θ based on $\mathcal{R}_\theta$, we find $C(\mathbf{s}, 1-\alpha)$ such that $P(\mathcal{R}_\theta(\mathbf{s}, \mathbb{S}^\star, \xi) < C(\mathbf{s}, 1-\alpha)) = 1-\alpha$, and the $100(1-\alpha)\%$ one-sided generalized upper confidence bound is taken to be $C(\mathbf{s}, 1-\alpha)$. This coincides exactly with the fiducial upper confidence bound provided that we construct the $\mathcal{R}_\theta(\mathbf{s}, \mathbb{S}^\star, \xi)$ in such a way that its distribution is the same as the fiducial distribution of θ. Our general construction for FGPQs shows that this is always possible. For this reason, we have singled out this subclass of GPQs and called them fiducial GPQs (FGPQs).

It is important to realize the strong connection that exists between GCIs and fiducial intervals, if for no other reason than to avoid reinventing interval procedures under a different name. For instance, some of the GCIs derived in the recent literature had already been derived using the fiducial argument during the 1950s and 1960s, or even earlier. In fact, the GCI for the difference between two normal means in the Behrens–Fisher problem turned out to coincide with the fiducial interval, as observed by Weerahandi (1991) himself. This comes as no surprise when the connection between GCIs and fiducial intervals is acknowledged. Likewise, Bross (1950) derived approximate percentiles for the fiducial distribution of σ_α^2 in the one-way random model (Example 2), and later Healy (1963) derived exact expressions for these percentiles. The resulting fiducial intervals coincide with the GCIs for σ_A^2 proposed by Weerahandi (1993), which again comes as no surprise.

Methods for constructing FGPQs discussed in Theorems 2 and 4 were motivated by discussions in the literature pertaining to development of fiducial distributions for parameters (see, e.g., Fisher 1935, 1939, 1970; Fraser 1961, 1966; Dawid and Stone 1982). In particular, Dawid and Stone (1982) provided a thorough investigation of the frequentist properties of fiducial intervals. Both Fraser (1961, 1966) and Dawid and Stone (1982) developed structural/fiducial distributions for parameters by conditioning on ancillary statistics when the dimension of the sufficient statistic exceeded the number of parameters. Their work implies that once a choice was made for the equations defining the manifold $\mathcal{M}(\mathbf{s})$ and an invertible pivotal relationship chosen (when one is available), a fiducial distribution for ξ, given a realized value $\mathbf{s}$ of $\mathbb{S}$, may be defined as the conditional distribution of $\mathbf{g}_0(\mathbf{s}, \mathbb{E})$ given that $\mathbb{B}(\mathbf{s}, \mathbb{E}) = \mathbf{0}$. As was the case for the construction of a FGPQ in Theorem 4, the fiducial distribution of ξ can, and often will, depend on the choice of the defining equations $\mathbb{B}(\mathbf{s}, \mathbb{E}) = \mathbf{0}$ for $\mathcal{M}(\mathbf{s})$.

It is not too difficult to see that once the invertible pivotal relationship (9) and the defining equations $\mathbb{B}(\mathbf{s}, \mathbb{E}) = \mathbf{0}$ for $\mathcal{M}(\mathbf{s})$ have been selected, there is a one-to-one correspondence between a FGPQ constructed according to Theorem 4 and the fiducial distribution obtained from the chosen invertible relationship by conditioning on $\mathbb{B}(\mathbf{s}, \mathbb{E})$. In particular, the distribution of $g(\mathbb{S}, \mathbf{G}(\mathbb{E}|\mathbf{0}))$ in Theorem 4 is the same as the fiducial distribution of ξ obtained by conditioning on $\mathbb{B}(\mathbf{s}, \mathbb{E})$ using (9). Thus GCIs obtained using Theorem 4 may also be obtained using the fiducial argument.

Throughout its long history, fiducial inference has been criticized on many grounds, one of which is the nonuniqueness of fiducial intervals in most problems. This nonuniqueness may arise from the particular invertible pivotal representation chosen or from the particular choice of an ancillary statistic or both. In this article we have demonstrated that the same nonuniqueness issues also arise for GCIs. What is noteworthy is that our main result regarding the asymptotic behavior of GCIs applies to any one of the several possible generalized pivots as long as the conditions of the theorem hold. We have given some examples that demonstrate this point.

We have also given an example where the conditions of the theorem do not hold. The parameter of interest in that example is the sum of squares of the means of several normal populations—an example previously considered by Wilkinson (1977) and Stein (1959) to demonstrate the *failure* of fiducial inference in certain problems. However, for this example, we propose an alternate construction of a GCI that has the correct asymptotic coverage.

Finally, although we have not elaborated on this, it is well known that many of the standard frequentist intervals, as well as fiducial intervals and GCIs, can be derived within a Bayesian framework with the appropriate choice of priors. This is not surprising, given the well-known complete-class results for Bayesian rules of inference. In such cases, asymptotic frequentist properties of Bayesian procedures may be invoked to demonstrate the asymptotic frequentist properties of GCIs and fiducial intervals. In specific instances, the challenge is to show that an appropriate prior exists that will lead to a particular GCI being considered. In some cases this is not even true. Grundy (1956) presented a class of one-parameter family of distributions for which the fiducial distribution of the parameter is not obtainable as the posterior distribution, no matter what prior is chosen. Hence there is some merit to demonstrating asymptotic frequentist properties of GCIs without relying on its Bayesian connections. It is noteworthy that Grundy's example satisfies the conditions of Theorem 1. This is no surprise, because it is *exact* by Theorem 5.

7. CONCLUDING REMARKS

In this article we have established an important result concerning GCIs introduced by Weerahandi (1993). Specifically, we have shown that under fairly mild conditions, GCIs have correct asymptotic frequentist coverage. To our knowledge, this is the first general result concerning the asymptotic behavior of GCIs. This result in turn implies corresponding asymptotic frequentist coverage properties for fiducial intervals by virtue of their connection with GCIs. This also appears to be the first result of this kind for fiducial intervals, which were originally proposed by Fisher (1935) and subsequently refined by Fraser (1966), who called them "structural intervals."

We have also provided general methods for constructing GPQs. The pivotal quantities derived by our constructions have certain additional useful properties; for instance, they automatically lead to generalized test variables and generalized tests. In addition, the distribution of the generalized pivot constructed by our recipes can be made to coincide with any fiducial distribution for the parameter in question. We feel that GPQs that have this property are appropriately termed FGPQs.

Although much of this article discusses the scalar parameter case, with a few appropriate additional assumptions, our theorem concerning the asymptotic behavior of generalized pivots extends to the vector case; see Appendix B. Although this article does not discuss tests in detail, it is worth emphasizing that Theorem 1 can be used in most common applications to verify that generalized tests for scalar parameters (Tsui and Weerahandi 1989) will have asymptotically correct type I error rates. Multivariate versions appear to be straightforward extensions of our work. A key issue that remains unresolved is the determination of a set of sufficient conditions under which the FGPQs obtained by the general construction will lead to confidence intervals with correct asymptotic coverage.

Finally, there is substantial overlap between the two concepts of fiducial inference and generalized inference. Nonuniqueness issues aside, for large classes of problems, both concepts allow development of inference procedures that have asymptotically correct frequentist behavior and small-sample performance that is often adequate for applications. FGPQs provide a framework for associating a distribution with a parameter in a general parametric setup.

APPENDIX A: PROOFS

Proof of Theorem 1

Define

$$H(\mathbf{s}) = \sum_{j=1}^{k} g_{1,j}(\mathbf{s},\xi) N_j. \tag{A.1}$$

Assumptions A.1 and A.2(b) imply that for all $\mathbf{s} \in \mathcal{A}$, $H(\mathbf{s})$ a is nondegenerate normal random variable. Denote its γ quantile by $C_H(\mathbf{s},\gamma)$.

We first prove that for all $\mathbf{s} \in \mathcal{A}$,

$$C_n(\mathbf{s},\gamma) = g_{0,n}(\mathbf{s},\xi) + \frac{C_H(\mathbf{s},\gamma)}{\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right). \tag{A.2}$$

Rewriting (2), we get

$$P_\xi\left(\sqrt{n}\big(\mathcal{R}_\theta(\mathbf{s},\mathbb{S}^\star,\xi) - g_{0,n}(\mathbf{s},\xi)\big) \le \sqrt{n}\big(C_n(\mathbf{s},\gamma) - g_{0,n}(\mathbf{s},\xi)\big)\right) = \gamma + o(1). \tag{A.3}$$

First, observe that Assumption A.2(c) implies that for all $\mathbf{s} \in \mathcal{A}$,

$$\sqrt{n} R_n(\mathbf{s},\mathbb{S}^\star,\xi) \xrightarrow{P_\xi} 0 \tag{A.4}$$

Next, observe that Slutsky's theorem, Assumption A.1, (1), (A.1), and (A.4) imply that

$$\sqrt{n}\big(\mathcal{R}_\theta(\mathbf{s},\mathbb{S}^\star,\xi) - g_{0,n}(\mathbf{s},\xi)\big) \xrightarrow{\mathcal{D}} H(\mathbf{s}). \tag{A.5}$$

By definition, $P_\xi(H(\mathbf{s}) \le C_H(\mathbf{s},\gamma)) = \gamma$. Therefore, by (A.3) and (A.5),

$$\sqrt{n}\big(C_n(\mathbf{s},\gamma) - g_{0,n}(\mathbf{s},\xi)\big) \to C_H(\mathbf{s},\gamma), \tag{A.6}$$

and (A.2) follows. Moreover, the continuity of the functions $g_{1,j}$ implies that $C_H(\mathbf{s},\gamma)$ is continuous at $\mathbf{s} = \mathbf{t}(\xi)$.

Now we can finish the proof. Combining (1), (A.2), (A.3), and (A.6), we have, for all $\varepsilon > 0$ small enough to guarantee $\{\mathbf{s} \mid \|\mathbf{s} - \mathbf{t}(\xi)\| < \epsilon\} \subset \mathcal{A}$,

$$P_\xi\Bigg(\|\mathbb{S} - \mathbf{t}(\xi)\| < \varepsilon, \sum_{j=1}^{k} g_{1,j,n}(\mathbb{S},\xi)\sqrt{n}\big(S_j - t_j(\xi)\big) + \sqrt{n}R_n(\mathbb{S},\mathbb{S},\xi) \le C_H(\mathbb{S},\gamma) + o(1)\Bigg)$$

$$\le P_\xi\big(\mathcal{R}_\theta(\mathbb{S},\mathbb{S},\xi) \le C_n(\mathbb{S},\gamma)\big)$$

$$\le P_\xi\Bigg(\|\mathbb{S} - \mathbf{t}(\xi)\| < \varepsilon, \sum_{j=1}^{k} g_{1,j,n}(\mathbb{S},\xi)\sqrt{n}\big(S_j - t_j(\xi)\big) + \sqrt{n}R_n(\mathbb{S},\mathbb{S},\xi) \le C_H(\mathbb{S},\gamma) + o(1)\Bigg) + P_\xi\big(\|\mathbb{S} - \mathbf{t}(\xi)\| \ge \varepsilon\big).$$

Notice that Assumption A.2(c) also implies that $\sqrt{n}R_n(\mathbb{S},\mathbb{S},\xi) \xrightarrow{P_\xi} 0$, and, by Slutsky's theorem, we get

$$\sum_{j=1}^{k} g_{1,j,n}(\mathbb{S},\xi)\sqrt{n}\big(S_j - t_j(\xi)\big) \xrightarrow{\mathcal{D}} H(\mathbf{t}(\xi)).$$

Thus, by the definition of convergence in distribution and probability, we observe, after some algebra, that $P_\xi(\mathcal{R}_\theta(\mathbb{S},\mathbb{S},\xi) \le C_n(\mathbb{S},\gamma)) \to P_\xi(H(\mathbf{t}(\xi)) \le C_H(\mathbf{t}(\xi),\gamma)) = \gamma$. This concludes the proof.

Proof of Proposition 2

Again, we need to verify the conditions of Theorem 1 for each fixed value of $\mu, \sigma_1^2, \ldots, \sigma_k^2$. First, set $n = \sum_{l=1}^{k} n_l$. It is well known that in this case,

$$\sqrt{n}\big(\bar{X}_1^\star - \mu, \ldots, \bar{X}_k^\star - \mu, S_1^{\star 2} - \sigma_1^2, \ldots, S_k^{\star 2} - \sigma_k^2\big) \xrightarrow{\mathcal{D}} (N_1, \ldots, N_{2k}),$$

where $N_1, \ldots, N_{2k}$ are independent Gaussian random variables. The generalized pivot may be written as

$$\mathcal{R}_\mu(\mathbb{S},\mathbb{S}^\star,\mu,\sigma_1^2,\ldots,\sigma_k^2) = \frac{\sum_{j=1}^{k} n_j \bar{X}_j / S_j^2}{\sum_{i=1}^{k} n_i / S_i^2} - \sum_{j=1}^{k} \frac{(n_j/S_j^2)(\bar{X}_j^\star - \mu)}{\sum_{i=1}^{k} n_i/S_i^2} + \sum_{j=1}^{k} \frac{(n_j/S_j^2)\sum_{l \ne j} n_l(\bar{X}_j - \bar{X}_l)/S_l^2}{(\sum_{i=1}^{k} n_i/S_i^2)^2 \sigma_j^2}\big(S_j^{\star 2} - \sigma_j^2\big) + R_n.$$

Assumption A requires that various conditions hold on an open neighborhood $\mathcal{A}$ of the true parameter $(\mu_1, \ldots, \mu_k, \sigma_1^2, \ldots, \sigma_k^2)$. To define this neighborhood, consider $0 < m < M$ such that $|\mu| < M$ and $m < \sigma_j^2 < M$ for all j, and define $\mathcal{A} = \{(\bar{X}_1, \ldots, \bar{X}_k, S_1^2, \ldots, S_k^2) \mid |\bar{X}_j| < M, m < S_j^2 < M, j = 1, \ldots, k\}$.

The exact calculations are standard, though tedious. To demonstrate the types of arguments required, we first show that

$$\frac{n_j/S_j^2}{\sum_{i=1}^{k} n_i/S_i^2} = \frac{n_j/(S_j^2 \sum_{l=1}^{k} n_l)}{\sum_{i=1}^{k} n_i/(S_i^2 \sum_{l=1}^{k} n_l)} \to \frac{c_j/S_j^2}{\sum_{i=1}^{k} c_i/S_i^2} \tag{A.7}$$

uniformly. Toward that end, observe that

$$\left|\sum_{i=1}^{k}\left(\frac{n_i}{S_i^2 \sum_{l=1}^{k} n_l} - \frac{c_i}{S_i^2}\right)\right| \le \frac{1}{m}\sum_{i=1}^{k}\left|\frac{n_i}{\sum_{l=1}^{k} n_l} - c_i\right| \to 0,$$

because both the numerator and the denominator converge uniformly for $(\bar{X}_1, \ldots, \bar{X}_k, S_1^2, \ldots, S_k^2) \in \mathcal{A}$. Also on $\mathcal{A}$, the limit of the denominator is uniformly bounded away from 0, and the limit of the numerator is uniformly bounded away from 0, the uniform convergence in (A.7) follows by standard arguments. Note also that in (A.7), the quantity in the limit is nonzero.

Similar, but somewhat more complicated calculations show that

$$\frac{(n_j/S_j^2)\sum_{l\neq j} n_l(\bar{X}_j-\bar{X}_l)/S_l^2}{(\sum_{i=1}^k n_i/S_i^2)^2\sigma_j^2} \to \frac{(c_j/S_j^2)\sum_{l\neq j} c_l(\bar{X}_j-\bar{X}_l)/S_l^2}{(\sum_{i=1}^k c_i/S_i^2)^2\sigma_j^2}$$

uniformly on $\mathcal{A}$. So $g_{1,j}$ satisfy conditions 2(a) and 2(b) of Assumption A. Finally, routine, though tedious, calculations show that the second partial derivatives of $\mathcal{R}_\mu(\mathbb{S},\mathbb{S}^\star,\mu,\sigma_1^2,\ldots,\sigma_k^2)$ are bounded on the neighborhood $\mathcal{A}$. The statement then follows from Theorem 1.

Proof of Proposition 3

Assume that $n_i/n \to c_i \in (0,1)$, where $n=\sum n_i$. We have $\sqrt{n}(\mathbb{S}_n - \mathbf{m}) \xrightarrow{\mathcal{D}} D\mathbb{Z}$, where $\mathbf{m}=(\mu,\ldots,\mu,\sigma_1^2,\ldots,\sigma_k^2)$, $\mathbb{Z}=(Z_1,\ldots,Z_{2k})$ are iid N(0, 1) variables and D is a diagonal matrix given by

$$D=\operatorname{diag}\left(\frac{\sigma_1}{\sqrt{c_1}},\ldots,\frac{\sigma_k}{\sqrt{c_k}},\frac{\sigma_1^2\sqrt{2}}{\sqrt{c_1}},\ldots,\frac{\sigma_k^2\sqrt{2}}{\sqrt{c_k}}\right).$$

By Skorokhod's theorem (see Billingsley 1995), we can find a sequence $\bar{\mathbb{S}}_n$ independent of $\mathbb{S}^*$ such that $\bar{\mathbb{S}}_n$ has the same distribution as $\mathbb{S}$ and $\sqrt{n}(\bar{\mathbb{S}}_n-\mathbf{m}) \to D\mathbb{Z}$ almost surely.

Close examination of the density in (10) shows that, conditionally on $\bar{\mathbb{S}}_n$,

$$\sqrt{n}\big(\mathcal{R}_\mu(\bar{\mathbb{S}}_n,\mathbb{S}^\star,\xi)-\mu\big) \to \mathrm{N}\left(\frac{\sum_{j=1}^k Z_j\sqrt{c_j}/\sigma_j}{\sum_{j=1}^k c_j/\sigma_j^2},\frac{1}{\sum_{j=1}^k c_j/\sigma_j^2}\right)\quad \text{a.s.}$$

If $C(\mathbf{s},n)$ is chosen so that $\lim_{n\to\infty} P_\xi(\mathcal{R}_\mu(\mathbf{s},\mathbb{S}^\star,\xi)\le C(\mathbf{s},n))=\alpha$, then we can prove in a similar way as in the proof of Theorem 1,

$$C(\bar{\mathbb{S}}_n,n)=\mu+\frac{1}{\sqrt{n}}\left(\frac{\sum_{j=1}^k Z_j\sqrt{c_j}/\sigma_j}{\sum_{j=1}^k c_j/\sigma_j^2}+\frac{z_\alpha}{(\sum_{j=1}^k c_j/\sigma_j^2)^{1/2}}\right) + o\left(\frac{1}{\sqrt{n}}\right), \quad (\text{A.8})$$

where z_α is the α quantile of the standard normal distribution. Finally, because the distribution of $\bar{\mathbb{S}}$ and $\mathbb{S}$ are the same, (A.8) implies that

$$\begin{aligned} P\big(\mu<C(\mathbb{S},n)\big) &= P\big(\mu<C(\bar{\mathbb{S}},n)\big)\\ &\to P\left(0<\frac{\sum_{j=1}^k Z_j\sqrt{c_j}/\sigma_j}{\sum_{j=1}^k c_j/\sigma_j^2}+\frac{z_\alpha}{(\sum_{j=1}^k c_j/\sigma_j^2)^{1/2}}\right)\\ &= \alpha. \end{aligned}$$

Therefore, just as in Theorem 1, we can conclude that the confidence intervals based on the FGPQ will have asymptotically correct coverage.

APPENDIX B: MULTIVARIATE VERSION OF THEOREM 1

In what follows we need the following notation.

Definition B.1. Open sets A_n converge to an open set A (i.e., $A_n \to A$) if $(\lim A_n)^\circ = A$. Here $\lim A_n = B$ exists if $I_{A_n} \to I_B$, I_A is the indicator function of A and B° is the interior of B.

Let us consider a parametric statistical problem where we observe $X_1,\ldots,X_n$, whose joint distribution belongs to some family of continuous distributions parameterized by $\xi\in\Xi\subset\mathbb{R}^p$. Let $\mathbb{S}=(S_1,\ldots,S_k)$ denote a statistic based on the X_i's. In theory, we can consider an independent copy of $X_1^\star,\ldots,X_n^\star$ and denote the statistic based on $X_i^\star$'s by $\mathbb{S}^\star$. Finally, suppose that a vector-valued function $\mathcal{R}_\theta(\mathbb{S},\mathbb{S}^\star,\xi)=(\mathcal{R}_{\theta,1}(\mathbb{S},\mathbb{S}^\star,\xi),\ldots,\mathcal{R}_{\theta,d}(\mathbb{S},\mathbb{S}^\star,\xi))$ is available that is a FGPQ for a parameter $\theta=\pi(\xi)\in\mathbb{R}^d$, $d\le k$.

In addition, assume that the following holds.

Assumption A′.

1. Assume that there exists $\mathbf{t}(\xi)\in\mathbb{R}^k$ such that

$$\sqrt{n}\big(S_1^\star - t_1(\xi),\ldots,S_k^\star-t_k(\xi)\big)\xrightarrow{\mathcal{D}}\mathbf{N}=(N_1,\ldots,N_k)^\top,$$

where $\mathbf{N}$ has a nondegenerate distribution multivariate normal distribution.

2. Assuming existence and continuity of second partial derivatives with respect to $\mathbf{s}^\star$ of $\mathcal{R}_{\theta,l}(\mathbf{s},\mathbf{s}^\star,\xi)$, we have the following one-term Taylor expansion with a remainder term:

$$\mathcal{R}_{\theta,l}(\mathbf{s},\mathbf{s}^\star,\xi)=g_{0,l,n}(\mathbf{s},\xi)+\sum_{j=1}^k g_{1,l,j,n}(\mathbf{s},\xi)\big(s_j^\star-t_j(\xi)\big) + R_{l,n}(\mathbf{s},\mathbf{s}^\star,\xi). \quad (\text{B.1})$$

Here

$$g_{0,l,n}(\mathbf{s},\xi)=\mathcal{R}_{\theta,l}\big(\mathbf{s},\mathbf{t}(\xi),\xi\big)\qquad\text{and}$$

$$g_{1,l,j,n}(\mathbf{s},\xi)=\frac{\partial}{\partial s_j^\star}\mathcal{R}_{\theta,l}(\mathbf{s},\mathbf{s}^\star,\xi)\bigg|_{\mathbf{s}^\star=\mathbf{t}(\xi)}.$$

Suppose that $A\subset\mathbb{R}^k$ is an open set containing $\mathbf{t}(\xi)$ with the following properties:

(a) The functions $g_{1,l,j,n}(\mathbf{s},\xi)$ converge uniformly in $\mathbf{s}\in A$ to a function $g_{1,l,j}(\mathbf{s},\xi)$ continuous at $\mathbf{s}=\mathbf{t}(\xi)$.

(b) The matrix

$$J(\mathbf{s})=\begin{pmatrix} g_{1,1,1}(\mathbf{s},\xi) & \cdots & g_{1,1,k}(\mathbf{s},\xi)\\ \vdots & \ddots & \vdots\\ g_{1,d,1}(\mathbf{s},\xi) & \cdots & g_{1,d,k}(\mathbf{s},\xi)\end{pmatrix}$$

is of rank d for all $\mathbf{s}\in A$.

(c) For each $l=1,\ldots,d$, the remainder $\sqrt{n}R_{l,n}(\mathbf{s},\mathbb{S}^\star,\xi)\xrightarrow{P_\xi}0$ uniformly in $\mathbf{s}$ on the open neighborhood A of $\mathbf{t}(\xi)$.

3. We consider a collection of open regions, $C(\mathbf{X},\gamma)\subset\mathbb{R}^d$ with $\lambda(\partial C(\mathbf{X},\gamma))=0$ (i.e., the boundary has zero Lebesgue measure), indexed by continuous random variables $\mathbf{X}$ and $\gamma\in(0,1)$ satisfying the following:

(a) $P(\mathbf{X}\in C(\mathbf{X},\gamma))=\gamma$.

(b) $C(a\mathbf{X}+b,\gamma)=aC(\mathbf{X},\gamma)+b$.

(c) If $\mathbf{X}$ has nondegenerate normal distribution, $\mathbf{X}_n\xrightarrow{\mathcal{D}}\mathbf{X}$, and $\gamma_n\to\gamma$, then $C(\mathbf{X}_n,\gamma_n)\to C(\mathbf{X},\gamma)$.

Recall that $R_\theta(\mathbb{S},\mathbb{S}^\star,\xi)$ has a distribution that is independent of the parameters. This allows us to state the following theorem.

Theorem B.1. Suppose that Assumption A′ holds and that $\gamma_n\to\gamma\in(0,1)$. Then

$$\lim_{n\to\infty}P_\xi\big(\theta=\mathcal{R}_\theta(\mathbb{S},\mathbb{S},\xi)\in C_n\big(R_\theta(\mathbb{S},\mathbb{S}^\star,\xi),\gamma_n\big)\big)=\gamma.$$

In particular, $C_n(R_\theta(\mathbb{S},\mathbb{S}^\star,\xi),\gamma)$ is a confidence region for θ with asymptotic coverage probability equal to γ.

Remark B.1. If $d=1$, then an example of $C(X,\gamma)=(-\infty,q(X,\gamma))$, where $q(X,\gamma)$ is the γ-quantile of the distribution of X. If $d>1$, then we can consider many different regions, for example, a cubical equal-tailed region.

Remark B.2. The various conditions stated in Assumption A′ could be weakened. For example, we do not have to assume that the limiting random variable $\mathbf{H}$ is normal. Assumption A′.3 and the proof of Theorem B.1 would then have to be modified accordingly.

[Received November 2004. Revised May 2005.]

REFERENCES

Barnard, G. A. (1977), "Pivotal Inference and the Bayesian Controversy," *Bulletin of the International Statistical Institute*, 47, 543–551.

——— (1981), "A Coherent View of Statistical Inference for Public Science," presented at Symposium on Statistical Inference and Applications, University of Waterloo, August 1981.

——— (1982), "Discussion: The Functional-Model Basis of Fiducial Inference," *The Annals of Statistics*, 10, 1068–1069.

——— (1995), "Pivotal Models and the Fiducial Argument," *International Statistical Review*, 63, 309–323.

Behrens, W.-V. (1929), "Ein Beitrag zur Fehlerberchnung bei Wenigen Beobachtungen," *Landwirtschaftliche Jahrbucher*, 68, 807–837.

Billingsley, P. (1995). *Probability and Measure* (3rd ed.), New York: Wiley.

Bross, I. (1950), "Fiducial Intervals for Variance Components," *Biometrics*, 6, 136–144.

Burdick, R. K., Borror, C. M., and Montgomery, D. C. (2005), *Design and Analysis of Gauge R&R Studies: Making Decisions With Confidence Intervals in Random and Mixed ANOVA Models*, Philadelphia: ASA/SIAM.

Burdick, R. K., and Park, D. J. (2003), "Performance of Confidence Intervals in Regression Models With Unbalanced One-Fold Nested Error Structures," *Communications in Statistics, Part B—Simulation and Computation*, 32, 717–732.

Burdick, R. K., Park, Y.-J., Montgomery, D. C., and Borror, C. M. (2005), "Confidence Intervals for Misclassification Rates in a Gauge R&R Study," *Journal of Quality Technology*, to appear.

Casella, G., and Berger, R. L. (2002), *Statistical Inference* (2nd ed.), Belmont, CA: Duxbury Press.

Chang, Y.-P., and Huang, W.-T. (2000), "Generalized Confidence Intervals for the Largest Value of Some Functions of Parameters Under Normality," *Statistica Sinica*, 10, 1369–1383.

Chiang, A. K. L. (2001), "Simple General Method for Constructing Confidence Intervals for Functions of Variance Components," *Technometrics*, 43, 356–367.

Chou, Y., and Owen, D. B (1984), "One-Sided Confidence Regions on the Upper and Lower Tail Areas of Normal Distribution," *Journal of Quality Technology*, 16, 150–158.

Cochran, W. G. (1964), "Approximate Significance Levels of the Behrens–Fisher Test," *Biometrics*, 20, 191–195.

Daniels, L., Burdick, R. K., and Quiroz, J. (2005), "Confidence Intervals in a Gauge R&R Study With Fixed Operators," *Journal of Quality Technology*, 37, 179–185.

Dawid, A. P., and Stone, M. (1982), "The Functional-Model Basis of Fiducial Inference," *The Annals of Statistics*, 10, 1054–1067.

Fisher, R. A. (1935), "The Fiducial Argument in Statistical Inference," *Annals of Eugenics*, VI, 91–98.

——— (1939), "Samples With Possibly Unequal Variances," *Annals of Eugenics*, IX, 174–180.

——— (1961a), "Sampling the Reference Set," *Sankhyā*, Ser. A, 23, 3–8.

——— (1961b), "The Weighted Mean of Two Normal Samples With Unknown Variance Ratio," *Sankhyā*, Ser. A, 23, 103–114.

——— (1970), *Statisitical Methods for Research Workers* (14th ed.), Oxford, U.K.: Oxford University Press.

Fraser, D. A. S. (1961), "On Fiducial Inference," *The Annals of Mathematical Statistics*, 32, 661–676.

——— (1966), "Structural Probability and a Generalization," *Biometrika*, 53, 1–9.

——— (1968), *The Structure of Inference*, New York: Krieger.

Graybill, F. A., and Wang, C. M. (1980), "Confidence Intervals on Nonnegative Linear Combinations of Variances," *Journal of the American Statistical Association*, 75, 869–873.

Grundy, P. M. (1956), "Fiducial Distributions and Prior Distributions: An Example in Which the Former Cannot Be Associated With the Latter," *Journal of the Royal Statistical Society*, Ser. B, 18, 217–221.

Hamada, M., and Weerahandi, S. (2000), "Measurement System Assessment via Generalized Inference," *Journal of Quality Technology*, 32, 241–253.

Hannig, J. (1996), "On Conditional Distributions as Limits of Martingales," unpublished dissertation, Charles University, Prague [in Czech].

——— (2005), "On Multidimensional Fiducial Generalized Confidence Intervals," Technical Report 2005/1, Colorado State University, Dept. of Statistics.

Healy, M. J. R. (1963), "Fiducial Limits for a Variance Component," *Journal of the Royal Statistical Society*, Ser. B, 25, 128–130.

Iyer, H. K., and Mathew, T. (2002), Comments on "Simple General Method for Constructing Confidence Interval for Functions of Variance Components" by A. K. L. Chiang, *Technometrics*, 44, 284–285.

Iyer, H. K., and Patterson, P. L. (2002), "A Recipe for Constructing Generalized Pivotal Quantities and Generalized Confidence Intervals," Technical Report 2002/10, Colorado State University, Dept. of Statistics.

Iyer, H. K., Wang, C. M., and Mathew, T. (2004), "Models and Confidence Intervals for True Values in Interlaboratory Trials," *Journal of the American Statistical Association*, 99, 1060–1071.

Johnson, N. L., and Kotz, S. (1970), *Continuous Univariate Distributions 2*, Boston: Houghton Mifflin.

Jordan, S. M., and Krishnamoorthy, K. (1996), "Exact Confidence Intervals for the Common Mean of Several Normal Populations," *Biometrics*, 52, 77–86.

Krishnamoorthy, K., and Lu, Y. (2003), "Inferences on the Common Mean of Several Normal Populations Based on the Generalized Variable Method," *Biometrics*, 59, 237–247.

Krishnamoorthy, K., and Mathew, T. (2003), "Inferences on the Means of i Lognormal Distributions Using Generalized p-Values and Generalized Confidence Intervals," *Journal of Statistical Planning and Inference*, 115, 103–121.

Linnik, J. V. (1968), *Statistical Problems With Nuisance Parameters*, translated from the 1966 Russian edition, Providence, RI: American Mathematical Society.

Mathew, T., and Krishnamoorthy, K. (2004), "One-Sided Tolerance Limits in Balanced and Unbalanced One-Way Random Models Based on Generalized Confidence Limits," Technometrics, 46, 44–52.

McNally, R. J., Iyer, H. K., and Mathew, T. (2001), "Tests for Individual and Population Bioequivalence Based on Generalized p Values," *Statistics in Medicine*, 22, 31–53.

Patterson, P. L., Hannig, J., and Iyer, H. K. (2004a), "Fiducial Generalized Confidence Intervals for the Common Mean of K Normal Populations," Technical Report 2004/10, Colorado State University, Dept. of Statistics.

——— (2004b), "Fiducial Generalized Confidence Intervals for Proportion of Conformance," Technical Report 2004/11, Colorado State University, Dept. of Statistics.

Peterson, J. J., Berger, V., and Weerahandi, S. (2003), "Generalized P-Values and Confidence Intervals: Their Role in Statistical Methods for Pharmaceutical Research and Development," technical report.

Rao, C. R. (1973), *Linear Statistical Inference and Its Applications*, New York: Wiley.

Roy, A., and Mathew, T. (2005), "A Generalized Confidence Limit for the Reliability Function of a Two-Parameter Exponential Distribution," *Journal of Statistical Planning and Inference*, 128, 509–517.

Satterthwaite, F. E. (1942), "A Generalized Analysis of Variance," *Annals of Mathematical Statistics*, 13, 34–41.

——— (1946), "An Approximate Distribution of Estimates of Variance Components," *Biometrics Bulletin*, 2, 110–114.

Stein, C. (1959), "An Example of Wide Discrepancy Between Fiducial and Confidence Intervals," *The Annals of Mathematical Statistics*, 30, 877–880.

Sukhatme, B. V. (1958), "Testing the Hypothesis That Two Populations Differ Only in Location," *The Annals of Mathematical Statistics*, 29, 60–78.

Tsui, K. W., and Weerahandi, S. (1989), "Generalized p-Values in Significance Testing of Hypotheses in the Presence of Nuisance Parameters," *Journal of the American Statistical Association*, 84, 602–607.

Tukey, J. W. (1951), "Components in Regression," *Biometrics*, 7, 33–69.

U.S. Food and Drug Administration (2001), "Guidance to Industry: Statistical Approaches to Establishing Bioequivalence," Rockville, MD: U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER).

Wang, C. M., and Lam, C. T. (1996), "Confidence Limits for Proportion of Conformance," *Journal of Quality Technology*, 28, 439–445.

Weerahandi, S. (1991), "Testing Variance Components in Mixed Models With Generalized p-Values," *Journal of the American Statistical Association*, 86, 151–153.

——— (1993), "Generalized Confidence Intervals," *Journal of the American Statistical Association*, 88, 899–905.

——— (1995), *Exact Statistical Methods for Data Analysis*, New York: Springer-Verlag.

——— (2004), *Generalized Inference in Repeated Measures*, New York: Wiley.

Welch, B. L. (1947), "The Generalization of 'Student's' Problem When Several Different Population Variances Are Involved," *Biometrika*, 34, 29–35.

Wilkinson, G. N. (1977), "On Resolving the Controversy in Statistical Inference," *Journal of the Royal Statistical Society*, Ser. B, 39, 119–171.

Williams, J. S. (1962), "A Confidence Interval for Variance Components," *Biometrics*, 49, 278–281.

Yu, P. L. H., Sun, Y., and Sinha, B. K. (1999), "On Exact Confidence Intervals for the Common Mean of Several Normal Populations," *Journal of Statistical Planning and Inference*, 81, 263–277.