

Comment

Jan Hannig*

University of North Carolina at Chapel Hill

Professors Mayo should be congratulated on bringing new light into the veritable arguments about statistical foundations. It is well documented that p-values, confidence intervals and hypotheses tests do not satisfy Strong Likelihood Principle. In the next section, we will demonstrate that fiducial distributions provide natural example of inference paradigm that breaks SLP while still satisfying Weak Conditionality Principle.

1. HISTORY OF FIDUCIAL INFERENCE

The origin of Generalized Fiducial Inference can be traced back to R. A. Fisher (Fisher, 1930, 1933, 1935) who introduced the concept of a fiducial distribution for a parameter, and proposed the use of this fiducial distribution, in place of the Bayesian posterior distribution. In the case of a one-parameter family of distributions, Fisher gave the following definition for a fiducial density $r(\theta)$ of the parameter based on a single observation x for the case where the cdf $F(x, \theta)$ is a decreasing function of θ :

$$(1.1) \quad r(\theta) = -\frac{\partial F(x, \theta)}{\partial \theta}.$$

For multiparameter families of distributions Fisher did not give a formal definition. Moreover, the fiducial approach led to confidence sets whose frequentist coverage probabilities were close to the claimed confidence levels but they were not exact in the frequentist sense. Fisher's proposal led to major discussions among the prominent statisticians of the 1930's, 40's and 50's (e.g., Dempster, 1966, 1968; Fraser, 1961b,a, 1966, 1968; Jeffreys, 1940; Lindley, 1958; Stevens, 1950). Many of these discussions focused on the nonexactness of the confidence sets and also nonuniqueness of fiducial distributions. The latter part of the 20th century has seen only a handful of publications Barnard (1995); Dawid *et al.* (1973); Salome (1998); Dawid and Stone (1982); Wilkinson (1977) as the fiducial approach fell into disfavor and became a topic of historical interest only.

Since the mid 2000s, there has been a true resurrection of interest in modern modifications of fiducial inference. These approaches have become known under the umbrella name of *distributional inference*. This increase of interest came both in the number of different approaches to the problem and the number of researchers working on these problems and manifested itself in an increasing number of publications in premier journals. The common thread for these approaches is a definition of inferentially meaningful probability statements about subsets of the parameter space without the need for subjective prior information.

330 Hanes Hall, Department of Statistics and OR, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3260 (e-mail: jan.hannig@unc.edu).

*Jan Hannig's research was supported in part by the National Science Foundation under Grant No. 1007543 and 1016441.

These modern approaches include Dempster-Shafer theory [Dempster \(2008\)](#); [Edlefsen et al. \(2009\)](#) and its recent extension called *inferential models* [Martin et al. \(2010\)](#); [Zhang and Liu \(2011\)](#); [Martin and Liu \(2013a,b,c,d\)](#). A somewhat different approach termed *confidence distributions* looks at the the problem of obtaining an inferentially meaningful distribution on the parameter space from purely frequentist point of view [Xie and Singh \(2013\)](#). One of the main contribution of this approach is the ability to combine information from disparate sources with deep implications for meta analysis [Schweder and Hjort \(2002\)](#); [Singh et al. \(2005\)](#); [Xie et al. \(2011\)](#); [Hannig and Xie \(2012\)](#); [Xie et al. \(2013\)](#). Another more mature approach is called *objective Bayesian inference* that aims at finding non-subjective model based priors. An example of a recent breakthrough in this area is the modern development of reference priors [Berger \(1992\)](#); [Berger and Sun \(2008\)](#); [Berger et al. \(2009, 2012\)](#); [Bayarri et al. \(2012\)](#). Another related approach is based on higher order likelihood expansions and implied data dependent priors [Fraser et al. \(2009\)](#); [Fraser \(2004, 2011\)](#); [Fraser and Naderi \(2008\)](#); [Fraser et al. \(2010, 2005\)](#). There is also important initial work showing how some simple fiducial distribution that are not Bayesian posteriors naturally arise within the decision theoretical framework [Taraldsen and Lindqvist \(2013\)](#).

Arguably, Generalized Fiducial Inference has been on the forefront of the modern fiducial revival. Starting in the early 1990s, the work of Tsui and Weerahandi [Tsui and Weerahandi \(1989, 1991\)](#) and Weerahandi [Weerahandi \(1993, 1994, 1995\)](#) on *generalized confidence intervals* and the work of [Chiang \(2001\)](#) on the *surrogate variable method* for obtaining confidence intervals for variance components, led to the realization that there was a connection between these new procedures and fiducial inference. This realization evolved through a series of works in the early 2000s [Hannig \(2009\)](#); [Hannig et al. \(2006\)](#); [Iyer et al. \(2004\)](#); [Patterson et al. \(2004\)](#). The strengths and limitations of the fiducial approach is becoming to be better understood, see, especially, [Hannig \(2009, 2013\)](#). In particular, the asymptotic exactness of fiducial confidence sets, under fairly general conditions, was established in [Hannig \(2013\)](#); [Hannig et al. \(2006\)](#); [Sonderegger and Hannig \(2013\)](#). Generalized fiducial inference has also been extended to prediction problems in [Wang et al. \(2012\)](#). Computational issues were discussed in [Cisewski and Hannig \(2012\)](#); [Hannig et al. \(2013\)](#) and model selection in the context of Generalized Fiducial Inference has been studied in [Hannig and Lee \(2009\)](#); [Lai et al. \(2013\)](#).

2. GENERALIZED FIDUCIAL DISTRIBUTION AND THE WEAK CONDITIONALITY PRINCIPLE

Most modern incarnations of fiducial inference begin with expressing the relationship between the data, \mathbf{X} , and the parameters, ξ , as

$$(2.1) \quad \mathbf{X} = \mathbf{G}(\mathbf{U}, \xi),$$

where $\mathbf{G}(\cdot, \cdot)$ is termed the *data generating equation* (also called the association equation or structural equation) and \mathbf{U} is the random component of this data generating equation whose distribution is free of parameters and completely known.

After observing the data \mathbf{x} the next step is to use the known distribution of \mathbf{U} and inverse of the data (2.1) to define probabilities for the subsets of the param-

eter space. In particular, Generalized Fiducial Inference defines a distribution on the parameter space as the weak limit as $\epsilon \rightarrow 0$ of the conditional distribution

$$(2.2) \quad \operatorname{argmin}_{\xi} \|\mathbf{x} - \mathbf{G}(\mathbf{U}^*, \xi)\| \mid \{\min_{\xi} \|\mathbf{x} - \mathbf{G}(\mathbf{U}^*, \xi)\| \leq \epsilon\},$$

where \mathbf{U}^* has the same distribution as \mathbf{U} . If there are multiple values minimizing the norm the operator $\operatorname{argmin}_{\xi}$ selects one of them (possibly at random). We stress at this point that Generalized Fiducial Distribution is not unique. For example different data generating equations can give somewhat different Generalized Fiducial Distribution. Notice also, that if $P(\min_{\xi} \|\mathbf{x} - \mathbf{G}(\mathbf{U}^*, \xi)\| = 0) > 0$, which is the case for discrete distributions, the limit in (2.2) is the conditional distribution evaluated at $\epsilon = 0$.

The conditional form of (2.2) implies the Weak Conditional Principle for the limiting Generalized Fiducial Distribution. To demonstrate this let us consider the two instrument example of Cox (1958) (see also section 4.1 of the discussed article). The data generating equation can be written in a hierarchical form:

$$\begin{aligned} M &= 1 + I_{(0,1/2)}(U) \\ X &= \theta + \sigma_M Z, \end{aligned}$$

where $U \sim U(0, 1)$ and $Z \sim N(0, 1)$ are independent and the precisions $\sigma_1 \ll \sigma_2$ are known. If both $X = x$ the measurement made and $M = m$ the instruments is used are observed, the conditional distribution (2.2) is $N(x, \sigma_m^2)$, only taking into account the experiment actually performed. On the other hand, if only M is unobserved then the conditional distribution (2.2) is the mixture $0.5N(\theta, \sigma_1^2) + 0.5N(\theta, \sigma_2^2)$. Generalized Fiducial Distribution therefore follows WCP as claimed.

3. GENERALIZED FIDUCIAL DISTRIBUTION AND THE STRONG LIKELIHOOD PRINCIPLE

In general, Generalized Fiducial Distribution does not satisfy the Strong Likelihood principle. We first demonstrate this on inference for geometric distribution. To begin we perform some preliminary calculations. Let X be a random variable with discrete distribution function $F(x, \xi)$. Let us assume for simplicity of presentation that for each fixed x , $F(x, \xi)$ is monotone in ξ and spans the whole $[0, 1]$. The inverse distribution function $F^{-1}(u, \xi) = \inf\{x : F(x, \xi) \geq u\}$ forms a natural data generating equation

$$X = F^{-1}(U, \xi), \quad U \sim (0, 1).$$

The minimizer in (2.2) is not unique but any fiducial distribution will have a distribution function satisfying $1 - F(x, \xi) \leq H(\xi) \leq 1 - F(x_-, \xi)$ if $F(x, \xi)$ is decreasing and $F(x_-, \xi) \leq H(\xi) \leq F(x, \xi)$ if $F(x, \xi)$ is increasing. To resolve this non-uniqueness Hannig (2009) and Efron (1998) recommend using the half correction which is the mixture distribution with distribution functions $H(\xi) = 1 - (F(x, \xi) + F(x_-, \xi))/2$ or $H(\xi) = (F(x, \xi) + F(x_-, \xi))/2$ respectively.

Let us now consider observing a random variable $N = n$ following Geometric(p) distribution. SLP implies that the inference based on observing $N = n$ should be the same as inference based on observing $X = 1$ where X is Binomial(n, p).

However the Geometric based Generalized Fiducial Distribution has a distribution function between $1 - (1 - p)^{n-1} \leq H_G(p) \leq 1 - (1 - p)^n$. The binomial based Generalized Fiducial Distribution uses bounds $1 - (1 - p)^n - np(1 - p)^{n-1} \leq H_B(p) \leq 1 - (1 - p)^n$. Thus the effect of the stopping rule demonstrate itself in the Generalized Fiducial Inference through the lower bound that is much closer to the upper bound in the case of geometric distribution. (We remark that one cannot ignore the lower bound as the upper bound is used to form upper confidence intervals and lower bound is used for lower confidence intervals on p .) To conclude, the fiducial distribution in this example depends on both the distribution function of x but also on the distribution function of $x - 1$.

Let us now turn our attention to continuous distributions. In particular, assume that the parameter $\xi \in \Theta \subset \mathbb{R}^p$ is p -dimensional and that the inverse to (2.1) $\mathbf{G}^{-1}(\mathbf{x}, \xi) = \mathbf{u}$ exists. Then under some differentiability assumptions, Hannig (2013) shows that the generalized fiducial distribution is absolutely continuous with density

$$(3.1) \quad r(\xi) = \frac{f(\mathbf{x}, \xi)J(\mathbf{x}, \xi)}{\int_{\Theta} f(\mathbf{x}, \xi')J(\mathbf{x}, \xi') d\xi'},$$

where $f(\mathbf{x}, \xi)$ is the likelihood and the function $J(\mathbf{x}, \xi)$ is

$$(3.2) \quad J(\mathbf{x}, \xi) = \sum_{\substack{\mathbf{i}=(i_1, \dots, i_p) \\ 1 \leq i_1 < \dots < i_p \leq n}} \left| \det \left(\frac{d}{d\xi} \mathbf{G}(\mathbf{u}, \xi) \Big|_{\mathbf{u}=\mathbf{G}^{-1}(\mathbf{x}, \xi)} \right)_{\mathbf{i}} \right|.$$

where $\frac{d}{d\xi} \mathbf{G}(\mathbf{u}, \xi)$ is the $n \times p$ Jacobian matrix of partial derivatives computed with respect of components of ξ . The sum in (3.2) spans over all p -tuples of indexes $\mathbf{i} = (1 \leq i_1 < \dots < i_p \leq n) \subset \{1, \dots, n\}$. Additionally, for any $n \times p$ matrix J , the sub-matrix $(J)_{\mathbf{i}}$ is the $p \times p$ matrix containing the rows $\mathbf{i} = (i_1, \dots, i_p)$ of A . The form of (3.1) suggests that as long as the Jacobian $J(\mathbf{x}, \xi)$ does not separate into $J(\mathbf{x}, \xi) = f(\mathbf{x})g(\xi)$, in which case the Generalized Fiducial Distribution is the same as Bayes posterior with $g(\xi)$ used as a prior, the Generalized Fiducial Distribution does not satisfy SLP due to the dependance on $d\mathbf{G}(\mathbf{u}, \xi)/d\xi$.

Heuristically, Generalized Fiducial Inference is using not only the data observed, but also the data that based on the data generating equation could have been observed in the neighborhood of the observed data.

4. GENERALIZED FIDUCIAL DISTRIBUTION AND SUFFICIENCY PRINCIPLE

Whether Generalized Fiducial Distribution satisfies sufficiency principle depends entirely on what data generating equations is chosen. For example, let us assume that $\mathbf{Y} = (\mathbf{S}(\mathbf{X}), \mathbf{A}(\mathbf{X}))'$, where \mathbf{S} is a p -dimensional sufficient and \mathbf{A} is ancillary and \mathbf{X} satisfies (2.1). Because $d\mathbf{A}/d\xi = 0$ the sum in (3.2) contains only one non-zero term:

$$(4.1) \quad J(\mathbf{x}, \xi) = \left| \det \left(\frac{d}{d\xi} \mathbf{S}(\mathbf{G}(\mathbf{u}, \xi)) \Big|_{\mathbf{u}=\mathbf{G}^{-1}(\mathbf{x}, \xi)} \right) \right|.$$

Let $\mathbf{s} = \mathbf{S}(\mathbf{x})$ and $\mathbf{a} = \mathbf{A}(\mathbf{x})$ be the observed values of the sufficient and ancillary statistics respectively. To interpret the Generalized Fiducial Distribution

assume that there is a unique ξ solving $\mathbf{s} = \mathbf{S}(\mathbf{G}(\mathbf{u}, \xi))$ for every \mathbf{u} and denote this solution $Q_{\mathbf{s}}(\mathbf{u}) = \xi$. Also assume that the ancillary data generating equation $\mathbf{A}(\mathbf{G}(\mathbf{u}, \xi)) = \mathbf{A}(\mathbf{u})$ is not a function of ξ . A straightforward calculation shows that the fiducial density (3.1) with (4.1) is the conditional distribution of $Q_{\mathbf{s}}(\mathbf{U}^*) \mid \mathbf{A}(\mathbf{U}^*) = \mathbf{a}$. We conclude that this choice of data generating equation leads to inference based on sufficient statistics conditional on the ancillary. However, we still do not expect the SLP to hold in general even for this data generating equation.

REFERENCES

- Barnard, G. A. (1995) Pivotal Models and the Fiducial Argument. *International Statistical Reviews*, **63**, 309–323.
- Bayarri, M. J., Berger, J. O., Forte, A. and Garca-Donato, G. (2012) Criteria for Bayesian Model Choice with Application to Variable Selection. *Annals of Statistics*, **40**, 1550 – 1577.
- Berger, J. O. (1992) On the development of reference priors [with discussion]. *Bayesian Statistics*, **4**, 35–60.
- Berger, J. O., Bernardo, J. M. and Sun, D. (2009) The formal definition of reference priors. *The Annals of Statistics*, **37**, 905–938.
- (2012) Objective Priors for Discrete Parameter Spaces. *Journal of the American Statistical Association*, **107**, 636–648.
- Berger, J. O. and Sun, D. (2008) Objective priors for the bivariate normal model. *The Annals of Statistics*, **36**, 963–982.
- Chiang, A. (2001) A simple general method for constructing confidence intervals for functions of variance components. *Technometrics. A Journal of Statistics for the Physical, Chemical and Engineering Sciences*, **43**, 356–367.
- Cisewski, J. and Hannig, J. (2012) Generalized fiducial inference for normal linear mixed models. *The Annals of Statistics*, **40**, 2102 – 2127.
- Cox, D. R. (1958) Some problems connected with statistical inference. *Ann. Math. Statist.*, **29**, 357–372.
- Dawid, A. P. and Stone, M. (1982) The functional-model basis of fiducial inference. *The Annals of Statistics*, **10**, 1054–1074.
- Dawid, A. P., Stone, M. and Zidek, J. V. (1973) Marginalization paradoxes in Bayesian and structural inference. *Journal of the Royal Statistical Society. Series B. Methodological*, **35**, 189–233.
- Dempster, A. P. (1966) New methods for reasoning towards posterior distributions based on sample data. *Ann. Math. Statist.*, **37**, 355–374.
- (1968) A generalization of Bayesian inference. (With discussion). *Journal of the Royal Statistical Society. Series B. Methodological*, **30**, 205–247.
- (2008) The Dempster-Shafer Calculus for Statisticians. *International Journal of Approximate Reasoning*, **48**, 365–377.
- Edlefsen, P. T., Liu, C. and Dempster, A. P. (2009) Estimating limits from Poisson counting data using Dempster–Shafer analysis. *The Annals of Applied Statistics*, **3**, 764–790.
- Efron, B. (1998) R.A. Fisher in the 21st century - Invited paper presented at the 1996 R.A. Fisher lecture. *Statistical Science. A Review Journal of the Institute of Mathematical Statistics*, **13**, 95–122.
- Fisher, R. A. (1930) Inverse Probability. *Proceedings of the Cambridge Philosophical Society*, **xxvi**, 528–535.
- (1933) The concepts of inverse probability and fiducial probability referring to unknown parameters. *Proceedings of the Royal Society of London series A*, **139**, 343–348.
- (1935) The Fiducial Argument in Statistical Inference. *Annals of Eugenics*, **VI**, 91–98.
- Fraser, A. M., Fraser, D. A. S. and Staicu, A.-M. (2009) The second order ancillary: A differential view with continuity. *Bernoulli. Official Journal of the Bernoulli Society for Mathematical Statistics and Probability*, **16**, 1208–1223.
- Fraser, D., Reid, N. and Wong, A. (2005) What a model with data says about theta. *Internat. J. Statist. Sci.*, **3**, 163–178.
- Fraser, D. A. S. (1961a) On fiducial inference. *Ann. Math. Statist.*, **32**, 661–676.
- (1961b) The fiducial method and invariance. *Biometrika*, **48**, 261–280.

- (1966) Structural probability and a generalization. *Biometrika*, **53**, 1–9.
- (1968) *The Structure of Inference*. New York-London-Sydney: John Wiley & Sons Inc.
- (2004) Ancillaries and Conditional Inference. *Statistical Science. A Review Journal of the Institute of Mathematical Statistics*, **19**, 333–369.
- (2011) Is Bayes posterior just quick and dirty confidence? *Statistical Science*, **26**, 299–316.
- Fraser, D. A. S. and Naderi, A. (2008) Exponential models: Approximations for probabilities. *Biometrika*, **94**, 1–9.
- Fraser, D. A. S., Reid, N., Marras, E. and Yi, G. Y. (2010) Default Priors for Bayesian and frequentist inference. *Journal of the Royal Statistical Society. Series B. Methodological*, **72**.
- Hannig, J. (2009) On Generalized Fiducial Inference. *Statistica Sinica*, **19**, 491–544.
- (2013) Generalized Fiducial Inference via Discretization. *Statistica Sinica*, **23**, 489 – 514.
- Hannig, J., Iyer, H. K. and Patterson, P. (2006) Fiducial generalized confidence intervals. *Journal of American Statistical Association*, **101**, 254 – 269.
- Hannig, J., Lai, R. C. S. and Lee, T. C. M. (2013) Computational issues of generalized fiducial inference. *Computational Statistics and Data Analysis*, **71**, 849–858.
- Hannig, J. and Lee, T. C. M. (2009) Generalized Fiducial Inference for Wavelet Regression. *Biometrika*, **96**, 847–860.
- Hannig, J. and Xie, M. (2012) A note on Dempster-Shafer Recombinations of Confidence Distributions. *Electronic Journal of Statistics*, **6**.
- Iyer, H. K., Wang, C. M. J. and Mathew, T. (2004) Models and confidence intervals for true values in interlaboratory trials. *Journal of the American Statistical Association*, **99**, 1060–1071.
- Jeffreys, H. (1940) Note on the Behrens-Fisher formula. *Ann. Eugenics*, **10**, 48–51.
- Lai, R. C. S., Hannig, J. and Lee, T. C. M. (2013) Generalized fiducial inference for ultra- high dimensional regression. Submitted for publication.
- Lindley, D. V. (1958) Fiducial distributions and Bayes’ theorem. *Journal of the Royal Statistical Society. Series B. Methodological*, **20**, 102–107.
- Martin, R. and Liu, C. (2013a) Conditional inferential models: combining information for prior-free probabilistic inference. Preprint.
- (2013b) Inferential models: A framework for prior-free posterior probabilistic inference. *Journal of the American Statistical Association*, **108**, 301 – 313.
- (2013c) Marginal inferential models: prior-free probabilistic inference on interest parameters. Preprint.
- (2013d) On a ‘plausible’ interpretation of p-values. Preprint.
- Martin, R., Zhang, J. and Liu, C. (2010) Dempster-Shafer theory and statistical inference with weak beliefs. *Statistical Science*, **25**, 72–87.
- Patterson, P., Hannig, J. and Iyer, H. K. (2004) Fiducial Generalized Confidence Intervals for Proportion of Conformance. *Tech. Rep. 2004/11*, Colorado State University.
- Salome, D. (1998) *Statistical Inference via Fiducial Methods*. Ph.D. thesis, University of Groningen.
- Schweder, T. and Hjort, N. L. (2002) Confidence and likelihood. *Scandinavian Journal of Statistics. Theory and Applications*, **29**, 309–332.
- Singh, K., Xie, M. and Strawderman, W. E. (2005) Combining information from independent sources through confidence distributions. *The Annals of Statistics*, **33**, 159–183.
- Sonderegger, D. and Hannig, J. (2013) Fiducial theory for free-knot splines. In *Festschrift in honor of Professor Hira L. Koul*. Springer.
- Stevens, W. L. (1950) Fiducial limits of the parameter of a discontinuous distribution. *Biometrika*, **37**, 117–129.
- Taraldsen, G. and Lindqvist, B. H. (2013) Fiducial theory and optimal inference. *Annals of Statistics*, **41**, 323 – 341.
- 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.
- (1991) Corrections: “Generalized p -values in significance testing of hypotheses in the presence of nuisance parameters” [J. Amer. Statist. Assoc. **84** (1989), no. 406, 602–607; MR1010352 (90g:62047)]. *Journal of the American Statistical Association*, **86**, 256.
- Wang, J. C.-M., Hannig, J. and Iyer, H. K. a. (2012) Fiducial Prediction Intervals. *Journal of Statistical Planning and Inference*, **142**, 1980–1990.
- Weerahandi, S. (1993) Generalized confidence intervals. *Journal of the American Statistical*

- Association*, **88**, 899–905.
- (1994) Correction: “Generalized confidence intervals” [J. Amer. Statist. Assoc. **88** (1993), no. **423**, 899–905; **MR1242940** (94e:62031)]. *Journal of the American Statistical Association*, **89**, 726.
- (1995) *Exact statistical methods for data analysis*. Springer Series in Statistics. New York: Springer-Verlag.
- Wilkinson, G. N. (1977) On resolving the controversy in statistical inference. *Journal of the Royal Statistical Society. Series B. Methodological*, **39**, 119–171.
- Xie, M., Liu, R. Y., Damaraju, C. V. and Olson, W. H. (2013) Incorporating external information in analyses of clinical trials with binary outcomes. *The Annals of Applied Statistics*, **7**, 342–368.
- Xie, M. and Singh, K. (2013) Confidence distribution, the frequentist distribution estimator of a parameter: A review. *International Statistical Review*, **81**.
- Xie, M., Singh, K. and Strawderman, W. E. (2011) Confidence Distributions and a Unifying Framework for Meta-Analysis. *Journal of the American Statistical Association*, **106**, 320–333.
- Zhang, J. and Liu, C. (2011) Dempster-Shafer inference with weak beliefs. *Statistica Sinica*, **21**, 475–494.