# A note on fiducial model averaging as an alternative to checking Bayesian and frequentist models

David R. Bickel

May 7, 2015

Ottawa Institute of Systems Biology
Department of Biochemistry, Microbiology, and Immunology
Department of Mathematics and Statistics
University of Ottawa
451 Smyth Road
Ottawa, Ontario, K1H 8M5

**Abstract**

Just as frequentist hypothesis tests have been developed to check model assumptions, prior predictive p values and other Bayesian p values check prior distributions as well as the model assumptions that yield the likelihood function. These model checks not only suffer from the usual threshold dependence of p values but also from suppressing model uncertainty in subsequent inference. The proposed solution is to transform Bayesian and frequentist p values for model assessment into a fiducial distribution across the models. Averaging the Bayesian or frequentist posterior distributions with respect to the fiducial distribution in some cases reproduces results from Bayesian model averaging or classical fiducial inference.

**Keywords:** Bayesian model averaging; coherent fiducial distribution; confidence distribution; distributional inference; fiducial inference; frequentist model averaging; model assessment; model checking; model criticism; posterior predictive check; prior-data conflict; prior predictive check

# 1 Introduction

A major motivation for model averaging from both Bayesian (Carlin and Louis, 2000, §2.4.2; Clyde and George, 2004, §1) and frequentist (Burnham and Anderson, 2002, §4.2.2; Claeskens and Hjort, 2008, §7.4) viewpoints is that uncertainty about the model is misleadingly suppressed when using the same data to select a model and then to make inferences on the basis of the selected model as if it were known *a priori*. In addition to that problem of the double-use of data, standard methods of model checking, including those with "Bayesian $p$ values" (Bernardo and Smith 1994, pp. 409-417; Bayarri and Berger 2004, §4.3; Little 2006; Gelman and Shalizi 2013, §4), rely on the arbitrary specification of a significance level or other threshold separating adequate models from inadequate models. Further, such methods only indicate whether a model is inadequate but do not measure the degree of inadequacy.

The purpose of this paper is to offer a framework for harnessing the potential of previous methods of model checking by overcoming all of the above problems. To accomplish this, frequency-valid $p$ values for model checking are transformed into fiducial distributions suitable for model averaging. Such $p$ values include not only those used to test the assumptions of frequentist models but also prior predictive $p$ values (Box, 1980) and some modified posterior predictive $p$ values (e.g., Hjort et al., 2006; Steinbakk and Storvik, 2009; Zhao and Xu, 2014) used to check priors and model assumptions that lead to likelihood functions. The proposed framework joins the ongoing renaissance in confidence distributions and other fiducial distributions (Nadarajah et al., 2015), extending their domain of application to territory currently held by Bayesian model averaging and frequentist model averaging (Wasserman, 2000; Burnham and Anderson, 2002; Hjort and Claeskens, 2003; Clyde and George, 2004; Claeskens and Hjort, 2008).

Fiducial model averaging applies to both Bayesian models, which include priors, and frequentist models, which do not. Just as Bayesian model averaging yields an average of Bayesian posterior distributions with respect to a Bayesian posterior distribution on the model space, a fiducial model average of Bayesian models yields an average of prior-based posterior distributions with respect to a fiducial distribution on the model space. A fiducial model average of frequentist models is the same as a fiducial average of Bayesian models except that the averaged posterior distributions based on priors are replaced with confidence distributions or other posterior distributions that do not require the specification of prior distributions.

This fiducial approach is widely applicable since a coherent fiducial distribution may be constructed from $p$ values arising in both frequentist and Bayesian model assessment. On the frequentist side, the $p$ value typically tests a small model embedded within a larger model, as when testing the hypothesis that a regression coefficient is zero for the purpose of covariate selection or the hypothesis that two variances are equal.

Fiducial averaging also applies to $p$ values used to test Bayesian models, including prior predictive $p$ values and posterior predictive $p$ values modified to be uniform on $[0, 1]$ (e.g., Box, 1980; Hjort et al., 2006; Steinbakk and Storvik, 2009; Zhao and Xu, 2014). Since such $p$ values are used to check models, Bickel (2015) framed fiducial averages of Bayesian models as a form of soft model checking. An empirical Bayes example of this averaging of Bayesian posterior distributions with respect to a fiducial distribution is the use of a confidence distribution to propagate model uncertainty to false discovery rates (Bickel, 2013, 2014). Fisher (1973, §5.6) had performed a similar operation in the context of "observations of two kinds."

Section 2 lays down the foundational concepts of models, confidence distributions, and fiducial distributions. Section 3 then defines fiducial model averaging and provides examples of averaging both Bayesian and frequentist models.

# 2 Preliminary notation and definitions

## 2.1 Frequentist models and Bayesian models

Without loss of generality, assume that the $n$-component vector observation $x^{\text{obs}}$ and the parameter $\psi$ of interest are continuous members of a sample space $\mathcal{X}$ and a parameter space $\Psi$, respectively. Cases of discrete or mixed $x^{\text{obs}}$ or $\psi$ may be easily obtained by replacing probability density functions with probability mass functions and integrals with sums.

The most common forms of frequentist model checking are frequentist model selection and testing models as null hypotheses (e.g., Burnham and Anderson, 2002; Claeskens and Hjort, 2008). In frequentist model checking, each model index $M \in \mathcal{M}$ corresponds to a family $\mathcal{F}_M = \{f_M(\bullet|\psi) : \psi \in \Psi\}$ of probability density functions on $\mathcal{X}$ but not to any prior distribution.

By contrast, in Bayesian model checking, a prior distribution is an additional component of the model (e.g., Bernardo and Smith, 1994, §6.1; Carlin and Louis, 2000, Ch. 6; Ando, 2010, §5.1). Given a set of models $\mathcal{M}$, the model indexed by $M \in \mathcal{M}$ specifies not only $\mathcal{F}_M$ but also a prior probability density function $\pi_M$ on $\Psi$. Thus, $f_M\left(x^{\text{obs}}|\psi\right)$ as a function of $\psi$ is an integrated likelihood function, with any nuisance parameters eliminated by marginalization with respect to their priors (see Berger et al., 1999). The posterior probability density of $\psi$ is $\pi_M\left(\psi|x^{\text{obs}}\right) = \pi_M\left(\psi\right) f_M\left(x^{\text{obs}}|\psi\right) / \int f_M\left(x^{\text{obs}}|\psi\right) \pi_M\left(\psi\right) d\psi$ according to model $M$.

In short, the index $M \in \mathcal{M}$ refers to $\mathcal{F}_M$, called a *frequentist model*, and, if priors are specified, to the pair $(\mathcal{F}_M, \pi_M)$, called a *Bayesian model*.

## 2.2 Fiducial distributions

### 2.2.1 Confidence probability distributions as basic fiducial distributions

The observation vector $x^{\text{obs}}$ is now considered as if it were a realization of $X \sim g_{\theta,\gamma}$, where $g_{\theta,\gamma}$ is a probability density function determined by a *basic parameter* $\theta$ and a *nonbasic parameter* $\gamma$, each named according to its relation to the following procedure (Bickel and Padilla, 2014). In this paper, the model $M$ and sometimes the interest parameter $\psi$ will be used as basic parameters.

Let $p^{x^{\text{obs}}}\left(\theta_0\right)$ denote an observed $p$ value for testing the null hypothesis that $\theta = \theta_0$ for all $\theta_0 \in \Theta$. The function $p^{\bullet}\left(\bullet\right) : \mathcal{X} \times \Theta \to [0,1]$ is a *confidence curve* if

$$\mathcal{K}^X\left(1-\alpha\right) = \left\{\theta \in \Theta : p^X\left(\theta\right) \geq \alpha\right\} \tag{1}$$

is a level-$(1-\alpha)$ confidence set, that is, if

$$\Pr\left(\theta \in \mathcal{K}^X\left(1-\alpha\right)|\theta\right) = 1-\alpha \tag{2}$$

for all $\alpha \in [0,1]$, $\theta \in \Theta$, and $\gamma \in \Gamma$ (Birnbaum, 1961; Blaker, 2000). Equation (2) holds if and only if $p^X\left(\theta\right) \sim \mathrm{U}\left(0,1\right)$ for all $\theta \in \Theta$ and $\gamma \in \Gamma$, where $\mathrm{U}\left(0,1\right)$ is the uniform distribution on $[0,1]$ (Bickel and Padilla, 2014). Let $\mathfrak{H}\left(x^{\text{obs}}\right)$ denote the set of all observed confidence sets: $\mathfrak{H}\left(x^{\text{obs}}\right) = \left\{\mathcal{K}^{x^{\text{obs}}}\left(1-\alpha\right) : \alpha \in [0,1]\right\}$. Let $\mathfrak{H}$ denote a $\sigma$-algebra of subsets of $\Theta$ such that $\mathfrak{H}\left(x^{\text{obs}}\right) \subset \mathfrak{H}$.

An additive measure $K^{x^{\text{obs}}}$ on the measurable space $(\Theta, \mathfrak{H})$ that assigns to each observed confidence set a mass equal to one of its confidence levels is known as a *confidence distribution*:

$$K^{x^{\text{obs}}}\left(\Theta_1\right) \in \left\{1-\alpha : \alpha \in [0,1], \mathcal{K}^{x^{\text{obs}}}\left(1-\alpha\right) = \Theta_1\right\} \tag{3}$$

for all $\Theta_1 \in \mathfrak{H}\left(x^{\text{obs}}\right)$. A confidence distribution is called a *basic fiducial distribution* if it has unit total mass $\left(K^{x^{\text{obs}}}\left(\Theta\right) = 1\right)$ (Bickel and Padilla, 2014). In that case, $K^{x^{\text{obs}}}$ is a Kolmogorov probability distribution with a random variable $\widetilde{\theta} \sim K^{x^{\text{obs}}}$.

The essential fiducial argument is captured as follows. The fiducial probability that $\widetilde{\theta}$ is in an observed confidence set is equal to the frequentist probability that $\theta$ is in a confidence set over repeated sampling:

**Lemma 1.** *Assume $\widetilde{\theta}$ is a random variable of basic fiducial distribution $K^{x^{\text{obs}}}$. For every $\alpha_1 \in [0,1]$, there is an $\alpha_2 \in [0,1]$ such that $\mathcal{K}^{x^{\text{obs}}}\left(1-\alpha_1\right) = \mathcal{K}^{x^{\text{obs}}}\left(1-\alpha_2\right)$ and*

$$\Pr\left(\widetilde{\theta} \in \mathcal{K}^{x^{\text{obs}}}\left(1-\alpha_1\right)\right) = \Pr\left(\theta \in \mathcal{K}^{X}\left(1-\alpha_2\right)\right).$$

*Proof.* Since $\widetilde{\theta} \sim K^{x^{\text{obs}}}$, equation (3) implies that

$$\begin{aligned}\Pr\left(\widetilde{\theta} \in \mathcal{K}^{x^{\text{obs}}}\left(1-\alpha_1\right)\right) &= K^{x^{\text{obs}}}\left(\mathcal{K}^{x^{\text{obs}}}\left(1-\alpha_1\right)\right) \\ &\in \left\{1-\alpha : \alpha \in [0,1], \mathcal{K}^{x^{\text{obs}}}\left(1-\alpha\right) = \mathcal{K}^{x^{\text{obs}}}\left(1-\alpha_1\right)\right\}.\end{aligned}$$

Thus, there is some $\alpha_2 \in [0,1]$ such that $1-\alpha_2 = \Pr\left(\widetilde{\theta} \in \mathcal{K}^{x^{\text{obs}}}\left(1-\alpha_1\right)\right)$ and $\mathcal{K}^{x^{\text{obs}}}\left(1-\alpha_2\right) = \mathcal{K}^{x^{\text{obs}}}\left(1-\alpha_1\right)$. Finally, $1-\alpha_2 = \Pr\left(\theta \in \mathcal{K}^{X}\left(1-\alpha_2\right)\right)$ is necessarily true by equation (2). □

While Wilkinson (1977) celebrated the "noncoherence" in confidence distributions satisfying $K^{x^{\text{obs}}}\left(\Theta\right) < 1$, the key condition for a confidence distribution to be a basic fiducial distribution $\left(K^{x^{\text{obs}}}\left(\Theta\right) = 1\right)$ is implicit in the treatments found in the confidence distribution literature reviewed by Nadarajah et al. (2015). For example, Fraser (1991), Schweder and Hjort (2002), Singh et al. (2005), Singh et al. (2007), and Bityukov et al. (2011) study the one-sided *p value function* or "significance function," the distribution function or survival function $p^{x^{\text{obs}}}\left(\bullet\right) : \Theta \to [0,1]$ of a basic fiducial distribution of scalar $\theta$. In the vector-$\theta$ case, certain generalized fiducial distributions (Hannig, 2009) and probability-matching Bayesian posterior distributions (Datta and Mukerjee, 2004) generate set estimates with approximate frequentist coverage (2) and thus qualify as approximate basic fiducial distributions. Other basic fiducial distributions with vector $\theta$ are mentioned in Bickel and Padilla (2014), which uses basic fiducial distributions in place of Bayesian posterior distributions for decision-theoretic purposes (Bickel, 2012a,b).

**Example 1.** The construction of a basic fiducial distribution from a one-sided $p$ value function is particularly straightforward for a scalar $\theta$. Suppose $\Theta \subseteq \mathbb{R}$ and that $p^{x^{\text{obs}}}(\theta_0)$ is a $p$ value for testing the null hypothesis that $\theta = \theta_0$ against the alternative hypothesis that $\theta > \theta_0$ such that $p^X(\theta_0) \sim \text{U}(0,1)$ for all $\theta_0 \in \Theta$, where $p^x(\bullet)$ is monotonically increasing for all $x \in \mathcal{X}$. Thus, $\mathcal{K}^x(1-\alpha) = ]\inf\Theta, \theta^x(\alpha)]$ for each $\alpha \in [0,1]$, where $\theta^x(\bullet)$ is the inverse of $p^x(\bullet)$ for all $x \in \mathcal{X}$, and the set of observed confidence intervals is $\mathfrak{H}\left(x^{\text{obs}}\right) = \left\{\left]\inf\Theta, \theta^{x^{\text{obs}}}(\alpha)\right] : \alpha \in [0,1]\right\}$. Let $\mathfrak{H}$ denote the set of Borel subsets of $\Theta$, and let $\widetilde{\theta}$ denote the random variable such that $p^{x^{\text{obs}}}(\bullet)$ is its distribution function, that is, $\Pr\left(\widetilde{\theta} \le \theta_0\right) = p^{x^{\text{obs}}}(\theta_0)$ for all $\theta_0 \in \Theta$. The probability distribution of $\widetilde{\theta}$ is denoted by $K^{x^{\text{obs}}}$ and satisfies the above requirements to be not only a confidence distribution but also a basic fiducial distribution. ▲

#### 2.2.2 Coherent fiducial distributions

As suggested by its name, a basic fiducial distribution is the basic building block for the construction of other fiducial distributions. A *coherent fiducial distribution* is recursively defined as probability measure on $(\Theta, \mathfrak{H})$ that is either a basic fiducial distribution or the distribution of a function of a random variable formed by marginalization or conditionalization of other coherent fiducial distributions (Bickel and Padilla, 2014). Unlike a basic fiducial distribution, a coherent fiducial distribution is not necessarily a confidence distribution. A *coherent fiducial density function* is a Radon-Nikodym derivative of a coherent fiducial distribution.

# 3 Fiducial model averaging

## 3.1 Definition of the fiducial model average

$\Pi^{x^{\text{obs}}}$ will denote a coherent fiducial distribution (§2.2.2) on $(\mathcal{M}, \mathfrak{M})$, where $\mathfrak{M}$ is a $\sigma$-field of subsets of $\mathcal{M}$. Each of the basic fiducial distributions on which $\Pi^{x^{\text{obs}}}$ is built requires the specification of a confidence curve $p^{\bullet}(\bullet)$ on $\mathcal{X} \times \mathcal{M}$, using $\mathcal{M}$ as the $\Theta$ in Section 2.2.1. That means that for every basic fiducial distribution, a $p$ value $p^{x^{\text{obs}}}(M)$ tests the model indexed by $M$ as a null hypothesis and that $p^X(M) \sim \mathrm{U}(0,1)$ for all $M \in \mathcal{M}$.

Let $\pi_M^{x^{\text{obs}}}$ denote the probability density function on $\Psi$ that depends also on the observation $x^{\text{obs}}$ and that obtains under the model of index $M \in \mathcal{M}$. Thus, each $\pi_M^{x^{\text{obs}}}(\psi)$ is the posterior probability density, such as a Bayesian posterior density (§3.3) or a frequentist posterior density (§3.4), according to model $M$. With respect to $\Pi^{x^{\text{obs}}}$, the *fiducial model average* of the posterior densities represented by $\mathcal{M}$ is the expectation value of the posterior probability density over the models sampled from $\Pi^{x^{\text{obs}}}$:

$$\bar{\pi}^{x^{\text{obs}}}(\psi) = \mathrm{E}_{\widetilde{M} \sim \Pi^{x^{\text{obs}}}}\left(\pi_{\widetilde{M}}^{x^{\text{obs}}}(\psi)\right) = \int_{\mathcal{M}} \pi_M^{x^{\text{obs}}}(\psi)\, d\Pi^{x^{\text{obs}}}(M) \tag{4}$$

for all $\psi \in \Psi$.

## 3.2 A special case for scalar-index models

A simple procedure for constructing a coherent fiducial distribution for use as $\Pi^{x^{\text{obs}}}$ is available if each $p$ value corresponds to at most two models, as is often the case for scalar $M$. For all $u \in [0,1]$, assuming that $\left|\left\{M \in \mathcal{M} : p^{x^{\text{obs}}}(M) = u\right\}\right| \in \{1,2\}$ and that $\mathcal{M} \subseteq \mathbb{R}$, let $M_0(u)$ and $M_1(u)$ denote the only members of $\mathcal{M}$ such that $p^{x^{\text{obs}}}(M_0(u)) = u$, $p^{x^{\text{obs}}}(M_1(u)) = u$, and $M_0(u) \le M_1(u)$, yielding the function $M_\bullet(\bullet) : \{0,1\} \times [0,1] \to \mathcal{M}$. The random variable $\widetilde{M}$ is defined by $\widetilde{M} = M_{\widetilde{B}}\left(\widetilde{U}\right)$ based on the independent random variables $\widetilde{U} \sim \mathrm{U}(0,1)$ and $\widetilde{B} \sim \mathrm{Bern}(1/2)$, the Bernoulli distribution with probability parameter $1/2$.

Let $K^{x^{\text{obs}}}$ denote the probability distribution of $\widetilde{M}$. Let $\mathcal{K}^x(1-\alpha) = \{M \in \mathcal{M} : p^x(M) \ge \alpha\}$ for all $x \in \mathcal{X}$ and $\alpha \in [0,1]$. Thus, $\mathcal{K}^X(1-\alpha)$ is a $(1-\alpha)100\%$ confidence interval for $M$ since $\Pr\left(M \in \mathcal{K}^X(1-\alpha)|M\right) = \Pr\left(p^X(M) \ge \alpha|M\right) = 1-\alpha$. In addition, equation (3) is satisfied since

$$\begin{aligned}
K^{x^{\text{obs}}}\left(\mathcal{K}^{x^{\text{obs}}}(1-\alpha)\right) &= \Pr_{\widetilde{M} \sim K^{x^{\text{obs}}}}\left(\widetilde{M} \in \mathcal{K}^{x^{\text{obs}}}(1-\alpha)\right) \\
&= \Pr_{\widetilde{U} \sim \mathrm{U}(0,1), \widetilde{B} \sim \mathrm{Bern}(1/2)}\left(p^x\left(M_{\widetilde{B}}\left(\widetilde{U}\right)\right) \ge \alpha\right) \\
&= \left(\Pr_{\widetilde{U} \sim \mathrm{U}(0,1)}\left(p^x\left(M_0\left(\widetilde{U}\right)\right) \ge \alpha|\widetilde{B} = 0\right) + \Pr_{\widetilde{U} \sim \mathrm{U}(0,1)}\left(p^x\left(M_1\left(\widetilde{U}\right)\right) \ge \alpha\right)|\widetilde{B} = 1\right)/2 \\
&= \left(\Pr_{\widetilde{U} \sim \mathrm{U}(0,1)}\left(\widetilde{U} \ge \alpha|\widetilde{B} = 0\right) + \Pr_{\widetilde{U} \sim \mathrm{U}(0,1)}\left(\widetilde{U} \ge \alpha\right)|\widetilde{B} = 1\right)/2 \\
&= \Pr_{\widetilde{U} \sim \mathrm{U}(0,1)}\left(\widetilde{U} \ge \alpha\right) = 1-\alpha
\end{aligned}$$

for all $\alpha \in [0,1]$. Therefore, $K^{x^{\text{obs}}}$ is a confidence distribution. Since $K^{x^{\text{obs}}}$ is also a probability distribution, it is a basic fiducial distribution, the simplest type of coherent fiducial distribution. This warrants setting $\Pi^{x^{\text{obs}}} = K^{x^{\text{obs}}}$ for the fiducial model average of equation (4), yielding

$$\bar{\pi}^{x^{\text{obs}}}(\psi) = \frac{1}{2}\int_0^1 \pi_{M_0(u)}^{x^{\text{obs}}}(\psi)\,du + \frac{1}{2}\int_0^1 \pi_{M_1(u)}^{x^{\text{obs}}}(\psi)\,du. \tag{5}$$

In the case that $M_0(u) = M_1(u)$ for all $u \in [0,1]$ and that $p^{x^{\text{obs}}}$ is a distribution function, as is usual if each $p^x(M)$ is a one-sided $p$ value,

$$\bar{\pi}^{x^{\text{obs}}}(\psi) = \int_0^1 \pi_{M(u)}^{x^{\text{obs}}}(\psi)\,du = \int_{\mathcal{M}} \pi_M^{x^{\text{obs}}}(\psi)\,\frac{dp^x(M)}{dM}dM, \tag{6}$$

where $M(u) = \left(p^{x^{\text{obs}}}\right)^{-1}(u)$.

### 3.3 Fiducial average of Bayesian models

In the case of averaging Bayesian models, the fiducial average $\bar{\pi}\left(\psi|x^{\text{obs}}\right)$ of the Bayesian models represented by $\mathcal{M}$ is given by equation (4) with the substitutions $\bar{\pi}^{x^{\text{obs}}}(\psi) = \bar{\pi}\left(\psi|x^{\text{obs}}\right)$ and $\pi_M^{x^{\text{obs}}}(\psi) = \pi_M\left(\psi|x^{\text{obs}}\right)$ for all $\psi \in \Psi$. Thus, fiducial averaging of Bayesian models is equivalent to standard Bayesian model averaging except that a posterior distribution over models is replaced with a fiducial distribution $\Pi^{x^{\text{obs}}}$, which does not require any prior distribution over the models.

The requirement in Section 2.2.1 that $p^X(M) \sim \text{U}(0,1)$ for all $M \in \mathcal{M}$ is met both for prior predictive $p$ values (Box, 1980) and for certain functions of posterior predictive $p$ values (Hjort et al., 2006; Steinbakk and Storvik, 2009; Zhao and Xu, 2014). Those Bayesian $p$ values rely on the concept of replicating data sets according to a prior and/or posterior. Under model $M$, the *prior replicated data vector* $X^{\text{prior}}$ and the *posterior replicated data vector* $X^{\text{post}}$ are the random variables distributed such that $X^{\text{prior}} \sim f_M$ and $X^{\text{post}} \sim f_M\left(\bullet|x^{\text{obs}}\right)$, according to the prior and posterior predictive density functions $f_M$ and $f_M\left(\bullet|x^{\text{obs}}\right)$, defined by $f_M(x) = \int f_M(x|\psi)\,\pi_M(\psi)\,d\psi$ and $f_M\left(x|x^{\text{obs}}\right) = \int f_M(x|\psi)\,\pi_M\left(\psi|x^{\text{obs}}\right)d\psi$ for all $x \in \mathcal{X}$, respectively. Some special cases appear in Sections 3.3.1-3.3.2.

### 3.3.1 Prior predictive $p$ values

Consider the $p$ value

$$p^{x}(M)=\Pr_{X^{\text{prior}}\sim f_M}\left(\tau_M\left(X^{\text{prior}}\right)\geq\tau_M(x)\right) \tag{7}$$

for all $x\in\mathcal{X}$, where the function $\tau_M(\bullet):\mathcal{X}\to\mathbb{R}$ transforms the data into a real statistic used to check the adequacy of each model of index $M\in\mathcal{M}$. The observed value $p^{x^{\text{obs}}}(M)$ is called a *prior predictive* $p$ value, which might overcome certain objections against significance testing (Box, 1980). The function $p^{\bullet}(\bullet)$ is a confidence curve since $p^{X^{\text{prior}}}(M)\sim\text{U}(0,1)$ by construction.

**Example 2.** The normal-normal model for $n$ observations with unknown mean $\psi$ and known variance $\sigma^2$ is specified by $\Theta=\mathbb{R}$, $x^{\text{obs}}=\left(x_1^{\text{obs}},\dots,x_n^{\text{obs}}\right)$, $x_i^{\text{obs}}\sim\text{N}\left(\psi,\sigma^2\right)$ for all $i=1,\dots,n$, and $\psi\sim\text{N}\left(\mu_0,\sigma_0^2\right)$ for some prior mean $\mu_0$ and prior variance $\sigma_0^2$. The posterior probability distribution of $\psi$ is $\text{N}\left(\mu_{\mu_0}\left(\bar{x}^{\text{obs}}\right),\sigma^2\left(\bar{x}^{\text{obs}}\right)\right)$, where $\mu_{\mu_0}\left(\bar{x}^{\text{obs}}\right)=\left(\mu_0+n\bar{x}^{\text{obs}}\sigma_0^2/\sigma^2\right)/\left(1+n\sigma_0^2/\sigma^2\right)$ and $\sigma^2\left(\bar{x}^{\text{obs}}\right)=\sigma_0^2/\left(1+n\sigma_0^2/\sigma^2\right)$.

The prior replicated sample mean $\bar{X}^{\text{prior}}=\sum_{i=1}^{n}X_i^{\text{prior}}/n$ has prior predictive distribution $\text{N}\left(\mu_0,\sigma^2/n+\sigma_0^2\right)$. Supposing that the value of $\mu_0$ but no other aspect of the model is subject to checking, $\mathcal{M}=\mathbb{R}$, and $\mu_0$ is the index of the model for each $\mu_0=M\in\mathbb{R}$. Let $\bar{x}$ denote the sample mean of $x$ for all $x\in\mathcal{X}=\mathbb{R}^n$. Setting $\tau_M(x)=\bar{x}$, equation (7) provides

$$\begin{aligned}p^{x^{\text{obs}}}(\mu_0) &= \Pr_{\bar{X}^{\text{prior}}\sim\text{N}\left(\mu_0,\sigma^2/n+\sigma_0^2\right)}\left(\bar{X}^{\text{prior}}\geq\bar{x}^{\text{obs}}\right)\\ &= 1-\Phi\left(\frac{\bar{x}^{\text{obs}}-\mu_0}{\sqrt{\sigma^2/n+\sigma_0^2}}\right)=\Phi\left(\frac{\mu_0-\bar{x}^{\text{obs}}}{\sqrt{\sigma^2/n+\sigma_0^2}}\right)\end{aligned}$$

as the prior-predictive $p$ value for each model of prior mean $\mu_0\in\mathbb{R}$, where $\Phi$ is the standard normal distribution function (Box, 1980). From Example 1, let $\Pi^{x^{\text{obs}}}=K^{x^{\text{obs}}}$ denote the basic fiducial distribution of distribution function $p^{x^{\text{obs}}}$. Thus, the fiducial random variable $\widetilde{\mu}_0$ representing the prior mean has fiducial distribution $\text{N}\left(\bar{x}^{\text{obs}},\sigma^2/n+\sigma_0^2\right)$, and the fiducial model average probability density of $\psi$ is

$$\bar{\pi}\left(\psi|x^{\text{obs}}\right)=\text{E}_{\widetilde{\mu}_0\sim\text{N}\left(\bar{x}^{\text{obs}},\sigma^2/n+\sigma_0^2\right)}\left(\pi_{\widetilde{\mu}_0}\left(\psi|x^{\text{obs}}\right)\right)=\int_{-\infty}^{\infty}\phi\left(\frac{\psi-\mu_{\mu_0}\left(\bar{x}^{\text{obs}}\right)}{\sigma\left(\bar{x}^{\text{obs}}\right)}\right)\phi\left(\frac{\mu_0-\bar{x}^{\text{obs}}}{\sqrt{\sigma^2/n+\sigma_0^2}}\right)d\mu_0 \tag{8}$$

according to equation (4), where $\phi$ is the standard normal density function. This agrees with the more generally applicable equation (6). Since $\Pi^{x^{\text{obs}}}$ is equal to the Bayesian posterior distribution from a flat prior over the models (the Lebesgue measure on the set of Borel subsets of $\mathbb{R}$), the fiducial model average (8) is identical to a Bayesian model average in this case. ▲

**Example 3.** The one-sided prior predictive $p$ value of Example 2 may be replaced by the two-sided prior predictive $p$ value:

$$p^{x^{\text{obs}}}(\mu_0) = \Pr_{\bar{X}^{\text{prior}} \sim \text{N}\left(\mu_0, \sigma^2/n+\sigma_0^2\right)} \left( \left(\bar{X}^{\text{prior}} - \mu_0\right)^2 \geq \left(\bar{x}^{\text{obs}} - \mu_0\right)^2 \right) = \Pr\left( \chi_1^2 \geq \frac{\left(\bar{x}^{\text{obs}} - \mu_0\right)^2 / \sigma^2}{1 + n\sigma_0^2/\sigma^2} n \right), \quad (9)$$

where $\chi_1^2$ is the $\chi^2$ random variable with 1 degree of freedom. Since $p^{x^{\text{obs}}}\left(\bar{x}^{\text{obs}} + \varepsilon\right) = p^{x^{\text{obs}}}\left(\bar{x}^{\text{obs}} - \varepsilon\right)$ for any $\varepsilon \geq 0$, this prior predictive $p$ value tests the null hypothesis that the prior density of $\psi$ is either $\phi\left(\frac{\psi-\mu_0}{\sigma_0}\right)$ or $\phi\left(\frac{\psi - \left(2\bar{x}^{\text{obs}} - \mu_0\right)}{\sigma_0}\right)$ for any $\mu_0 \geq \bar{x}^{\text{obs}}$. Applying the method of Section 3.2, equation (5) reduces to

$$\bar{\pi}\left(\psi | x^{\text{obs}}\right) = \Pr\left(\widetilde{B} = 0\right) \int_0^1 \phi\left( \frac{\psi - \mu_{-\mu_0(u)}\left(\bar{x}^{\text{obs}}\right)}{\sigma\left(\bar{x}^{\text{obs}}\right)} \right) du + \Pr\left(\widetilde{B} = 1\right) \int_0^1 \phi\left( \frac{\psi - \mu_{\mu_0(u)}\left(\bar{x}^{\text{obs}}\right)}{\sigma\left(\bar{x}^{\text{obs}}\right)} \right) du, \quad (10)$$

where $\mu_0(u)$ is the nonnegative solution of $p^{x^{\text{obs}}}(\mu_0(u)) = u$ and $\Pr\left(\widetilde{B} = 0\right) = \Pr\left(\widetilde{B} = 1\right) = 1/2$.

Alternatively, two fiducial distributions may be combined as follows. Instead of assuming $\widetilde{B} \sim \text{Bern}(1/2)$, the parameter of the Bernoulli distribution determining the sign may be equated with the probability that $\widetilde{\mu}_0 > 0$, using the fiducial random variable $\widetilde{\mu}_0$ defined in Example 2. That is,

$$\widetilde{B} \sim \text{Bern}\left( \Pr_{\widetilde{\mu}_0 \sim \text{N}\left(\bar{x}^{\text{obs}}, \sigma^2/n+\sigma_0^2\right)} \left(\widetilde{\mu}_0 > 0\right) \right) = \text{Bern}\left( 1 - \Phi\left( -\frac{\bar{x}^{\text{obs}}}{\sqrt{\sigma^2/n + \sigma_0^2}} \right) \right).$$

The fiducial model average is still expressed as equation (10) but now with $\Pr\left(\widetilde{B} = 0\right) = \Phi\left(-\frac{\bar{x}^{\text{obs}}}{\sqrt{\sigma^2/n+\sigma_0^2}}\right)$ and $\Pr\left(\widetilde{B} = 1\right) = 1 - \Pr\left(\widetilde{B} = 0\right)$. This alternative approach is a special case of the development of Wilkinson (1977) appearing in Bickel and Padilla (2014, §6). Therein, the distribution of $\left(2\widetilde{B} - 1\right)\sqrt{\widetilde{\theta}}$ is presented as a coherent fiducial distribution, where $\widetilde{\theta}$ is the random variable whose distribution function is the $p^{x^{\text{obs}}}(\bullet)$ of equation (9). Thus, $2\widetilde{B} - 1$ and $\sqrt{\widetilde{\theta}}$ are fiducial random variables targeting the sign and magnitude of $\mu_0$, respectively. While each of the random variables $\widetilde{B}$ and $\widetilde{\theta}$ is constructed from a basic fiducial distribution, the coherent fiducial distribution of $\left(2\widetilde{B} - 1, \sqrt{\widetilde{\theta}}\right)$ is not a basic fiducial distribution since it is not a joint confidence distribution for the sign and magnitude of $\mu_0$. ▲

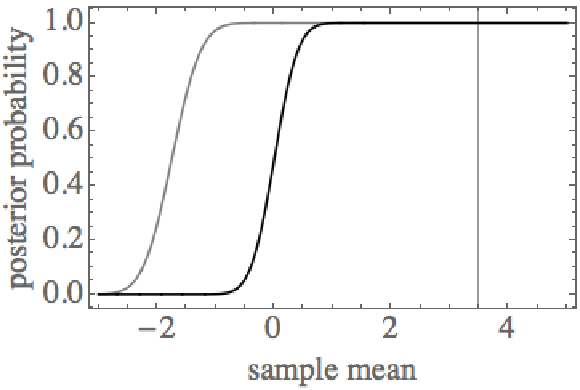


Figure 1: The posterior probabilities that $\psi > 0$ versus $\bar{x}^{\text{obs}}$. Black: $\int_0^\infty \bar{\pi}\left(\psi|x^{\text{obs}}\right) d\psi$ with $\bar{\pi}\left(\psi|x^{\text{obs}}\right)$ from equation (8). Gray: $\int_0^\infty \phi\left(\left(\psi - \mu_{3.5}\left(\bar{x}^{\text{obs}}\right)\right)/\sigma\left(\bar{x}^{\text{obs}}\right)\right) d\psi$, the probability under the model corresponding to the vertical line $\left(\bar{x}^{\text{obs}} = \mu_0 = 3.5\right)$. That model is assessed in Figure 2.

Figures 1-2 display the fiducial model average from Example 2 and the prior predictive $p$ value from Example 3 using these settings from Walter and Augustin (2009) and Bickel (2015): $n = 10$, $\sigma^2 = 1$, $\sigma_0^2 = 1/5$, and, for the model to be checked, $\mu_0 = 3.5$. A comparison of the plots reveals that the same conclusion would be reached about the sign of $\psi$, whether using the initial model ($\mu_0 = 3.5$) or the model average, when $\bar{x}^{\text{obs}}$ is such that the model is considered adequate $\left(p^{x^{\text{obs}}}(3.5) \geq 0.05\right)$. On the other hand, when the $\mu_0 = 3.5$ model is deemed inadequate $\left(p^{x^{\text{obs}}}(3.5) < 0.05\right)$, only the inference from fiducial model averaging remains unambiguous.

#### 3.3.2 Posterior predictive $p$ values

With $X^{\text{post}}$ drawn from the posterior distribution rather than the prior distribution and with a function $\tau_M(\bullet,\bullet) : \mathcal{X} \times \Psi \to \mathbb{R}$ in place of $\tau_M(\bullet)$, equation (7) becomes

$$p_{\text{post}}^{x}(M) = \Pr_{X^{\text{post}} \sim f_M(\bullet|x), \psi \sim \pi_M(\bullet|x)}\left(\tau\left(X^{\text{post}}, \psi\right) \geq \tau(x, \psi)\right),$$

for all $x \in \mathcal{X}$ and $M \in \mathcal{M}$. The function $\tau_M(\bullet,\bullet)$ in the model-checking literature is a measure of discrepancy that, unlike a statistic, may depend on $\psi$ (Hjort et al., 2006). The *posterior predictive* $p$ value of the model of index $M$ is the observed value, $p_{\text{post}}^{x^{\text{obs}}}(M)$.

Since the distribution of $p_{\text{post}}^{X^{\text{prior}}}(M)$, typically being very different from $\mathrm{U}(0,1)$, leads to problems with

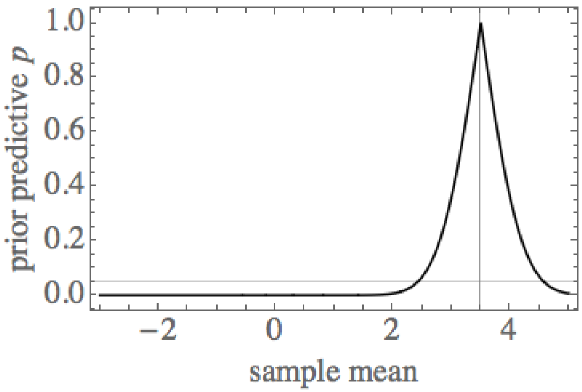


Figure 2: The prior predictive $p$ value $p^{x^{\text{obs}}}(3.5)$ versus $\bar{x}^{\text{obs}}$ according to equation (9). The vertical line is at $\bar{x}^{\text{obs}} = \mu_0 = 3.5$, representing the model of the gray line in Figure 1, and the horizontal line is the 0.05 significance threshold.

interpreting $p_{\text{post}}^{x^{\text{obs}}}(M)$, adjustments have been proposed. For instance, Hjort et al. (2006) suggested the *calibrated posterior predictive* $p$ value,

$$p_{\text{cpost}}^{x}(M) = \Pr_{X^{\text{prior}} \sim f_M}\left(p_{\text{post}}^{X^{\text{prior}}}(M) \leq p_{\text{post}}^{x}(M)\right). \tag{11}$$

Since $p_{\text{post}}^{x^{\text{obs}}}(M)$ is a statistic (Hjort et al., 2006, §10.3), $p_{\text{cpost}}^{x^{\text{obs}}}(M)$ is a prior predictive $p$ value in disguise: the substitution $\tau_M(\bullet) = -p_{\text{post}}^{\bullet}(M)$ in equation (7) yields $p_{\text{cpost}}^{x^{\text{obs}}}(M) = p^{x^{\text{obs}}}(M)$, resulting in $p_{\text{cpost}}^{X^{\text{prior}}}(M) \sim \text{U}(0,1)$. Thus, since $p_{\text{cpost}}^{\bullet}$ meets the requirements of a confidence curve (1), $p_{\text{cpost}}^{x^{\text{obs}}}(M)$ is suitable for constructing coherent fiducial distributions even though $p_{\text{post}}^{x^{\text{obs}}}(M)$ alone is not.

**Example 4.** For the set of models used in Examples 2 and 3, Hjort et al. (2006, (12)) reported

$$p_{\text{cpost}}^{x^{\text{obs}}}(\mu_0) = \Pr\left(\chi_1^2 \geq \frac{\left(\bar{x}^{\text{obs}} - \mu_0\right)^2/\sigma^2}{1 + n\sigma_0^2/\sigma^2} n\right).$$

Since equation (9) indicates that $p_{\text{cpost}}^{x^{\text{obs}}}(\mu_0) = p^{x^{\text{obs}}}(\mu_0)$, the calibrated posterior predictive $p$ value is equal to the prior predictive $p$ value for this set of models. Reasoning analogous to that of Examples 2 and 3 again yields the fiducial model averages given by equations (8) and (10), respectively. ▲

### 3.4 Fiducial average of frequentist models

In the case of averaging frequentist models, $\pi_M^{x^{\text{obs}}}(\psi)$ may be a coherent fiducial density (§2.2.2), a bootstrap density (Kroese and Schaafsma, 2004), a Dempster-Shafer density (Dempster, 2008; Martin et al., 2010), or any other posterior probability density that does not depend on a prior distribution. Without a prior, the methods based on Bayesian $p$ values (§3.3) are not applicable, but analogous methods based on a non-Bayesian $p$ value of each model leads to confidence distributions that generate coherent fiducial densities for use as $\pi_M^{x^{\text{obs}}}(\psi)$.

In the special case that each $\pi_M^{x^{\text{obs}}}$ of equation (4) is a coherent fiducial density function (§2.2), the fiducial model average is the density function corresponding to the *marginal fiducial distribution* defined by Bickel and Padilla (2014). Thus, the usual fiducial distributions that are marginal fiducial distributions are also fiducial model averages. In the following example, fiducial model averaging agrees with Bayesian model averaging using certain default priors.

**Example 5.** In the Behrens-Fisher problem (Cohen and Kim, 2014), $X_{1,j} \sim \text{N}\left(\mu_1, \sigma_1^2\right)$ for $j = 1, \ldots, n_1$ and $X_{2,j} \sim \text{N}\left(\mu_2, \sigma_2^2\right)$ for $j = 1, \ldots, n_2$, with all $X_{i,j}$ independent $(i = 1, 2)$, all parameters except the sample sizes $n_1$ and $n_2$ unknown, and the mean difference $\psi = \mu_1 - \mu_2$ is of interest. The observable vectors are $\boldsymbol{X}_1 = (X_{1,1}, \ldots, X_{1,n_1})$ and $\boldsymbol{X}_2 = (X_{2,1}, \ldots, X_{2,n_2})$, and the observed values of $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are $\boldsymbol{x}_1^{\text{obs}}$ and $\boldsymbol{x}_2^{\text{obs}}$, respectively. The sample means and standard deviations are denoted by $\hat{\mu}\left(\boldsymbol{x}_1^{\text{obs}}\right)$, $\hat{\mu}\left(\boldsymbol{x}_2^{\text{obs}}\right)$, $\hat{\sigma}\left(\boldsymbol{x}_1^{\text{obs}}\right)$, and $\hat{\sigma}\left(\boldsymbol{x}_2^{\text{obs}}\right)$. Let $x^{\text{obs}} = \left(\boldsymbol{x}_1^{\text{obs}}, \boldsymbol{x}_2^{\text{obs}}\right)$, $\widetilde{\mu}_1 = T_1 \hat{\sigma}\left(\boldsymbol{x}_1^{\text{obs}}\right)/\sqrt{n_1} + \hat{\mu}\left(\boldsymbol{x}_1^{\text{obs}}\right)$, and, $\widetilde{\mu}_2 = T_2 \hat{\sigma}\left(\boldsymbol{x}_2^{\text{obs}}\right)/\sqrt{n_2} + \hat{\mu}\left(\boldsymbol{x}_2^{\text{obs}}\right)$, where $T_1$ and $T_2$ are independent draws from the Student $t$ distribution with $n_1 - 1$ and $n_2 - 1$ degrees of freedom. The probability distributions of $\widetilde{\mu}_1$ and $\widetilde{\mu}_2$, denoted respectively by $K^{\boldsymbol{x}_1^{\text{obs}}}$ and $K^{\boldsymbol{x}_2^{\text{obs}}}$, are basic fiducial distributions since they are confidence distributions for $\mu_1$ and $\mu_2$. Let $\kappa^{\boldsymbol{x}_1^{\text{obs}}}$ and $\kappa^{\boldsymbol{x}_2^{\text{obs}}}$ represent the corresponding density functions with respect to the Lebesgue measure.

In the notation of Section 2.1, the set of frequentist models is $\mathcal{F}_{\mu_2} = \{f_{\mu_2}(\bullet|\psi) : \psi \in \mathbb{R}\}$ for all $\mu_2 \in \mathbb{R}$, where $f_{\mu_2}(\bullet|\psi)$ is the joint probability density function of $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$. For the model of each index $\mu_2$, let $\pi_{\mu_2}^{\boldsymbol{x}_1^{\text{obs}}}$ denote the probability density function of the random variable $\widetilde{\psi}_{\mu_2} = \widetilde{\mu}_1 - \mu_2$ with respect to the

Lebesgue measure. With $\Pi^{x^{\text{obs}}} = K_2^{x^{\text{obs}}}$, equation (4) yields

$$\begin{aligned}\bar{\pi}^{x^{\text{obs}}}(\psi) &= \mathrm{E}_{\widetilde{\mu}_2 \sim K_2^{x^{\text{obs}}}}\left(\pi_{\widetilde{\mu}_2}^{x^{\text{obs}}}(\psi)\right) = \int_{\mathbb{R}} \pi_{\mu_2}^{\boldsymbol{x}_1^{\text{obs}}}(\psi)\,\kappa^{\boldsymbol{x}_2^{\text{obs}}}(\mu_2)\,d\mu_2 \\ &= \int_{-\infty}^{\infty} \kappa^{\boldsymbol{x}_1^{\text{obs}}}(\psi - (-\mu_2))\,\kappa^{\boldsymbol{x}_2^{\text{obs}}}(\mu_2)\,d\mu_2,\end{aligned}$$

the convolution of the density functions of $\widetilde{\mu}_1$ and $-\widetilde{\mu}_2$. Thus, $\widetilde{\psi}_{\widetilde{\mu}_2} \sim \bar{\pi}^{x^{\text{obs}}}$ given the independence of $\widetilde{\mu}_1$ and $-\widetilde{\mu}_2$, and $\bar{\pi}^{x^{\text{obs}}}$ is the density function corresponding to the Behrens-Fisher fiducial distribution (Fisher, 1935), which is equal to the Bayesian posterior density of $\mu_1 - \mu_2$ obtained from the joint prior density of $(\mu_1, \mu_2, \sigma_1, \sigma_2)$ proportional to $\sigma_1^{-1}\sigma_2^{-1}$ (Jeffreys, 1940). While $\bar{\pi}^{x^{\text{obs}}}$ is a coherent fiducial density function, its interval estimates are too conservative for it to qualify as a basic fiducial distribution or other confidence distribution (Bickel and Padilla, 2014, Exa. 6). ▲

## Acknowledgments

This research was partially supported by the Natural Sciences and Engineering Research Council of Canada, by Agriculture and Agri-Food Canada, and by the Faculty of Medicine of the University of Ottawa.

## References

Ando, T., 2010. Bayesian Model Selection and Statistical Modeling. Statistics: A Series of Textbooks and Monographs. Taylor & Francis.

Bayarri, M. J., Berger, J. O., 2004. The interplay of Bayesian and frequentist analysis. Statistical Science 19, 58–80.

Berger, J. O., Liseo, B., Wolpert, R. L., 1999. Integrated likelihood methods for eliminating nuisance parameters. Statistical Science 14, 1–28.

Bernardo, J. M., Smith, A. F. M., 1994. Bayesian Theory. John Wiley and Sons, New York.

Bickel, D. R., 2012a. Coherent frequentism: A decision theory based on confidence sets. Communications in Statistics - Theory and Methods 41, 1478–1496.

Bickel, D. R., 2012b. Controlling the degree of caution in statistical inference with the Bayesian and frequentist approaches as opposite extremes. Electron. J. Statist. 6, 686–709.

Bickel, D. R., 2013. Simple estimators of false discovery rates given as few as one or two p-values without strong parametric assumptions. Statistical Applications in Genetics and Molecular Biology 12, 529–543.

Bickel, D. R., 2014. A fiducial continuum from confidence sets to empirical Bayes set estimates as the number of comparisons increases. Working Paper, University of Ottawa, deposited in uO Research at http://hdl.handle.net/10393/31898.

Bickel, D. R., 2015. Inference after checking multiple Bayesian models for data conflict. In revision.

Bickel, D. R., Padilla, M., 2014. A prior-free framework of coherent inference and its derivation of simple shrinkage estimators. Journal of Statistical Planning and Inference 145, 204–221.

Birnbaum, A., 1961. Confidence curves: An omnibus technique for estimation and testing statistical hypotheses. Journal of the American Statistical Association 56, pp. 246–249.

Bityukov, S., Krasnikov, N., Nadarajah, S., Smirnova, V., 2011. Confidence distributions in statistical inference. AIP Conference Proceedings 1305, 346–353.

Blaker, H., 2000. Confidence curves and improved exact confidence intervals for discrete distributions. Canadian Journal of Statistics 28 (4), 783–798.

Box, G., 1980. Sampling and Bayes' inference in scientific modelling and robustness. Journal of the Royal Statistical Society A 143, 383–430.

Burnham, K. P., Anderson, D. R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer, New York.

Carlin, B. P., Louis, T. A., 2000. Bayes and Empirical Bayes Methods for Data Analysis, 2nd Edition. Chapman & Hall/CRC, New York.

Claeskens, G., Hjort, N. L., 2008. Model Selection and Model Averaging. Cambridge University Press, Cambridge.

Clyde, M., George, E. I., 2004. Model uncertainty. Statistical Science 19, 81–94.

Cohen, A. S., Kim, S.-H., 2014. Behrens-fisher problem. In: Lovric, M. (Ed.), International Encyclopedia of Statistical Science. Springer Berlin Heidelberg, pp. 138–141.

Datta, G. S., Mukerjee, R., 2004. Probability Matching Priors: Higher Order Asymptotics. Springer, New York.

Dempster, A. P., 2008. The Dempster-Shafer calculus for statisticians. International Journal of Approximate Reasoning 48, 365–377.

Fisher, R., 1935. The fiducial argument in statistical inference. Annals of Human Genetics 6, 391–398.

Fisher, R. A., 1973. Statistical Methods and Scientific Inference. Hafner Press, New York.

Fraser, D. A. S., 1991. Statistical inference: likelihood to significance. Journal of the American Statistical Association 86, 258–265.

Gelman, A., Shalizi, C. R., 2013. Philosophy and the practice of Bayesian statistics. British Journal of Mathematical & Statistical Psychology 66 (1), 8–38.

Hannig, J., 2009. On generalized fiducial inference. Statistica Sinica 19, 491–544.

Hjort, N., Dahl, F., Steinbakk, G., 2006. Post-processing posterior predictive p values. Journal of the American Statistical Association 101 (475), 1157–1174.

Hjort, N. L., Claeskens, G., 2003. Frequentist model average estimators. Journal of the American Statistical Association 98, 879–899.

Jeffreys, H., 1940. Note on the behrens-fisher formula. Annals of Eugenics 10 (1), 48–51.

Kroese, A. H., Schaafsma, W., 2004. Distributional inference. In: Encyclopedia of Statistical Sciences. John Wiley and Sons, Inc.

Little, R. J., 2006. Calibrated Bayes: a Bayes/frequentist roadmap. The American Statistician 60 (3), 213–223.

Martin, R., Zhang, J., Liu, C., 2010. Dempster-Shafer theory and statistical inference with weak beliefs. Statistical Science 25 (1), pp. 72–87.

Nadarajah, S., Bityukov, S., Krasnikov, N., 2015. Confidence distributions: A review. Statistical Methodology 22, 23–46.

Schweder, T., Hjort, N. L., 2002. Confidence and likelihood. Scandinavian Journal of Statistics 29, 309–332.

Singh, K., Xie, M., Strawderman, W. E., 2005. Combining information from independent sources through confidence distributions. Annals of Statistics 33, 159–183.

Singh, K., Xie, M., Strawderman, W. E., 2007. Confidence distribution (CD) – distribution estimator of a parameter. IMS Lecture Notes Monograph Series 2007 54, 132–150.

Steinbakk, G. H., Storvik, G. O., 2009. Posterior Predictive p-values in Bayesian Hierarchical Models. Scandinavian Journal of Statistics 36 (2), 320–336.

Walter, G., Augustin, T., 2009. Imprecision and prior-data conflict in generalized bayesian inference. Journal of Statistical Theory and Practice 3 (1), 255–271.

Wasserman, L., 2000. Bayesian model selection and model averaging. Journal of Mathematical Psychology 44 (1), 92–107.

Wilkinson, G. N., 1977. On resolving the controversy in statistical inference (with discussion). Journal of the Royal Statistical Society B 39, 119–171.

Zhao, G., Xu, X., 2014. The one-sided posterior predictive p-value for fieller's problem. Statistics and Probability Letters 95, 57–62.