Fiducial Distributions for Real Exponential Families

Eugenio Melilli and Piero Veronese

Abstract The fiducial argument was introduced by Fisher in order to obtain distributions for unknown parameters without the need of a bayesian perspective. In recent years, a certain interest has grown for fiducial inference. In this paper we are using a result obtained by Petrone and Veronese in order to construct a fiducial distribution for the parameter of a discrete or continuous real exponential family in a simple and quite general manner. We identify the families for which a fiducial distribution can be seen as a posterior with respect to a (improper) prior, thus completing previous results by Lindley and we demonstrate that such a prior belongs to the conjugate family. Some further results on the fiducial distribution are discussed.

Key words: asymptotics, Bayesian posterior distribution, conjugate prior, coverage probability, simple quadratic variance function

1 Introduction and Preliminaries

Fiducial inference, after being introduced by Fisher in the 1930s (see, for example, [3]), had a significant growth in subsequent years but was substantially abandoned later due to its difficulties in interpretation and deficiencies when trying to extend it to less elementary models, see [7]. The main aim of Fisher's fiducial inference was to transfer randomness from the observed quantity to the parameter, in order to build a probability distribution for the parameter capturing all the information given by the data, without the need of a bayesian prior. This *fiducial distribution* could then be used in a straightforward way to derive inferences on the parameter, mainly to obtain confidence intervals. Originally Fisher considered a one-parameter con-

Eugenio Melilli
Bocconi University, Milan, e-mail: eugenio.melilli@unibocconi.it

Piero Veronese
Bocconi University, Milan, e-mail: piero.veronese@unibocconi.it

tinuous statistical model for which the distribution function (df) $F_\theta(x)$ is a strictly decreasing and differentiable function of the parameter $\theta \in \Theta$. For each observed x, he defined the fiducial density of θ as the function $h_x(\theta) = -\partial F_\theta(x)/\partial\theta$, provided that $\lim_{\theta\to\sup\Theta} F_\theta(x) = 1 - \lim_{\theta\to\inf\Theta} F_\theta(x) = 0$. More recently a resurgence of interest for fiducial inference took place. A very interesting and comprehensive discussion on generalized fiducial inference is given by Hannig [4].

In this paper we consider a *Natural Exponential Family* (NEF) with density, with respect to a non-degenerate σ-finite measure ν, given by $p_\theta(x) = \exp\{\theta x - M(\theta)\}$, $\theta \in \Theta$, where $M(\theta) = \ln\int \exp\{\theta x\}\nu(dx)$ and Θ is the interior of the *natural parameter* space $\{\theta \in \mathbb{R} : M(\theta) < \infty\}$. It is well known that, for an i.i.d. sample $(X_1,\dots,X_n)$ from a NEF, $S = \sum_{i=1}^n X_i$ is a sufficient statistic having df $F_{n,\theta}$ and density, with respect to a suitable measure ν_n,

$$p_{n,\theta}(s) = \exp\{\theta s - nM(\theta)\}, \quad \theta \in \Theta.$$

In Section 2 we define a fiducial distribution for the parameter of a continuous or discrete NEF. Moreover, we establish a relationship between expectations and maximum likelihood estimates, the asymptotic normality of the fiducial distributions and we discuss the frequentist coverage of fiducial intervals. In Section 3 we characterize fiducial distributions that can be seen as posteriors, together with the corresponding priors, completing a previous result by Lindley [5].

2 Fiducial distributions in continuous and discrete NEF's

The starting point of our analysis is a result by Petrone and Veronese [6] which essentially establishes that, if S is distributed according to $F_{n,\theta}$, then $H_{n,s}(\theta) \equiv 1 - F_{n,\theta}(s)$ is a continuous df on Θ. Hence its density

$$h_{n,s}(\theta) = \frac{\partial}{\partial\theta} H_{n,s}(\theta) = -\frac{\partial}{\partial\theta} F_{n,\theta}(s) = \int_{(s,+\infty)} (t - nM'(\theta)) p_{n,\theta}(t) d\nu_n(t) \quad (1)$$

can be interpreted as the fiducial density of the natural parameter θ. From (1), the fiducial distribution of a general parameter $\lambda = g(\theta)$ can be trivially obtained through the standard change-of-variable rule. The function $h_{n,s}$ in (1) is a density for both continuous and discrete NEF's. However, while for continuous NEF's replacing $F_{n,\theta}$ with its left-continuous version has no effects, in the discrete case this is not true. Two different versions of the fiducial distribution of θ are obtained and we denote by $H^L_{n,s}$ the one generated from the left-continuous version of $F_{n,\theta}$. This non-uniqueness in the discrete case typically arises in all approaches to fiducial inference; see e.g. [4]. Moreover, both $H_{n,s}$ and $H^L_{n,s}$ can fail to be df's for some boundary values of s. For instance, when $F_{n,\theta}$ is the binomial df with success probability $p = \exp(\theta)/(1+\exp(\theta))$, $F_{n,\theta}(n) = 1$ and thus $H_{n,n}(\theta) = 0$ for all θ, so that $H_{n,s}$ is not a df for $s = n$. Similarly $H^L_{n,s}$ is not a df for $s = 0$. We will comment on this point in Section 3.

The following examples provide the fiducial distribution computed making use of (1) for some common statistical models.

- *Normal model with mean μ and known variance σ^2*. The fiducial distribution of μ is Normal with mean $\bar{x} = s/n$ and variance σ^2/n.
- *Negative Exponential model with mean λ^{-1}*. The fiducial distribution of λ is Gamma with parameters s and n.
- *Bernoulli model with success probability p*. The fiducial distribution of p is Beta$(s+1, n-s)$ for the right-continuous version of $F_{n,\theta}$ and Beta$(s, n-s+1)$ for the left-continuous one.
- *Poisson model with mean λ*. The fiducial distribution of λ is Gamma$(n, s+1)$ for the first version and Gamma$(n+1, s)$ for the second one.
- *Geometric model with success probability p*. The fiducial distribution of p is Beta$(n, s+1)$ for the first version and Beta$(n+1, s)$ for the second one.

Some interesting results can be achieved using the connection between $H_{n,s}(\theta)$ and $F_{n,\theta}(s)$,

Proposition 1. *Given a continuous NEF with natural parameter θ and quadratic variance function (i.e., the variance is quadratic in the mean parameter $\mu = M'(\theta) = E_\theta(S/n)$), the expected value of θ, with respect to $H_{n,s}$, is*

$$E^{H_{n,s}}(\theta) = M'^{-1}(s/n) = M'^{-1}(\bar{x}).$$

Thus $M'^{-1}(\bar{x})$ can be seen as a fiducial estimate of θ. Since $\bar{x}$ is the maximum likelihood estimate (MLE) $\hat{\mu}$ of $\mu = M'(\theta)$ and, for the MLE invariance property, $\hat{\theta} = M'^{-1}(\bar{x})$, then the fiducial estimate of θ coincides with $\hat{\theta}$. As a consequence, the fiducial estimate of a general parameter $\lambda = g(\theta)$, with g smooth, is equal to $\hat{\lambda} = g(\hat{\theta})$, because of the definition of the fiducial distribution of λ.

The following useful asymptotic result can be easily proven.

Proposition 2. *Under the fiducial distribution, the mean parameter μ of a NEF is asymptotically normal with mean $\bar{x} = s/n$ and variance $M''(M'^{-1}(s/n))/n$.*

When using fiducial distributions for statistical inference one usually has, following Fisher's spirit, a frequentist perspective. Hence, as observed by Hannig [4], it is important to ensure that fiducial distributions lead to procedures that are (at least approximately) exact in the frequentist sense. This should be specifically required for confidence sets. Then, if $q_\alpha(s)$ denotes the $(1-\alpha)$-quantile of $H_{n,s}$, i.e. $H_{n,s}(q_\alpha(s)) = 1-\alpha$, one would like to have $Pr(\{s : \theta \le q_\alpha(s)\}) \simeq 1-\alpha$, where the last probability is computed under $F_{n,\theta}$. This is of course analogous to the widely discussed problem of matching priors in a bayesian setting, see [2]. For continuous NEF's the previous requirement is satisfied in an "exact" sense, while for the discrete ones the validity of the result is only approximate, i.e. the frequentist coverage of the set $\{s : \theta < q_\alpha(s)\}$ converges to $1-\alpha$ for $n \to +\infty$. However, the behaviors of $H_{n,s}$ and $H^L_{n,s}$ are quite different for small n. An interesting discussion about the coverage of intervals for a binomial proportion is given in [1].

3 Connections between fiducial and posterior distributions

The existence of a prior generating a posterior equal to the fiducial distribution in the case of a continuous NEF was discussed by Lindley [5]. He showed that such a prior exists only for Gaussian and Gamma models. The following proposition generalizes this result and gives a complete answer to the problem.

Proposition 3. *The fiducial distribution $H_{n,s}$ for a real NEF coincides with a posterior only for the following families having quadratic variance function: Normal (known variance), Binomial, Poisson, Gamma (known shape) and Negative Binomial. The corresponding prior for the natural parameter θ is proportional to $M'(\theta)$ in all cases except for the Normal one, in which the prior is constant.*

Moreover, $H_{n,s}$ belongs to the usual conjugate family whose density is $\pi_\theta(\theta|s',n') \propto \exp\{\theta s' - n'M(\theta)\}$, with $s' = s+l$ and $n' = n-q$, where q and l are the coefficients of the quadratic variance function $V(\mu) = q\mu^2 + l\mu + c$ of the NEF.

A similar conclusion holds for discrete families if we consider $H^L_{n,s}$; in this case, the "updating rule" is $s' = s+l-1$ and $n' = n-q$. Recognizing that a fiducial distribution belongs to the conjugate family is relevant since it allows directly to use its well known good mathematical and statistical properties.

Finally we observe that, since there is no argument to favor $H_{n,s}$ or $H^L_{n,s}$ as fiducial distribution in the discrete case, one could take a sort of "mean" $H^M_{n,s}$ between them. If we take $H^M_{n,s}$ as the conjugate with parameter $s' = s+l-1/2$, given by the average of $s+l$ and $s+l-1$, then it is easy to see that it turns out to be the posterior corresponding to the Jeffreys' prior. This proposal solves the non-uniqueness problem yielding a fiducial distribution always defined for all observations and with good coverage properties, as proved in [1] and [4].

Acknowledgements This research was supported by a grant of "L. Bocconi" University.

References

1. Brown, L.D.(2001). Interval estimation for a binomial proportion (with comment). Statistical Science **16**,101-133.
2. Datta, G. S. and Mukerjee, R. (2004). Probability Matching Priors: Higher Order Asymptotics. Lecture Notes in Statistics. Springer, New York.
3. Fisher, R.A. (1973). Statistical Methods and Scientific Inference, 3rd ed. Hafner Press, New York.
4. Hannig J. (2009). On generalized fiducial inference. Statistica Sinica. **19**, 491-544.
5. Lindley, D.,V. (1958). Fiducial distributions and Bayes' Theorem, J. Roy. Statist. Soc. Ser. B **20**, 102-107.
6. Petrone, S. and Veronese, P. (2010). Feller operators and mixture priors in Bayesian nonparametrics. Statistica Sinica. **20**, 379-404.
7. Zabell, S.L. (1992). R.A. Fisher and the Fiducial Argument Statist. Sci. **7**, 369-387.