

# THE POSTERIOR DISTRIBUTION OF THE LIKELIHOOD RATIO AS A MEASURE OF EVIDENCE

I. Smith and A. Ferrari

*Laboratoire H. Fizeau, UNS, CNRS, OCA  
(isabelle.smith@unice.fr, ferrari@unice.fr)*

**Abstract.** This paper deals with simple versus composite hypothesis testing under Bayesian and frequentist settings. The Posterior distribution of the Likelihood Ratio (PLR) concept is proposed in [1] for significance testing. The PLR is shown to be equal to 1 minus the p-value in a simple case. The PLR is used in [2] in order to calibrate p-values, Fractional Bayes Factors (FBF) and others. Dempster's equivalence result is slightly extended by adding a nuisance parameter in the test.

On the other hand, in [3] the p-values and the posterior probability of the null hypothesis  $\Pr(H_0|x)$  (seen as a Bayesian measure of evidence against the null hypothesis) are shown to be irreconcilable. Actually, as emphasized in [4],  $\Pr(H_0|x)$  is a measure of accuracy of a test, not a measure of evidence in a formal sense because it does not involve the likelihood ratio.

The PLR may give such a measure of evidence and be related to a natural p-value. In this presentation, in a classical invariance framework the PLR with inner threshold 1 will be shown to be equal to 1 minus a p-value where the test statistics is the likelihood, weighted by a term that accounts for some volume distortion effect.

Other analytical properties of the PLR will be proved in more general settings. The minimum of its support is equal to the Generalized Likelihood Ratio if  $H_0$  is nested in  $H_1$  and its moments are directly related to the (F)BF for a proper prior. Its relation to credible domains is also studied.

Practical issues will also be considered. The PLR can be implemented using a simple Monte Carlo Markov Chain and will be applied to extrasolar planet detection using direct imaging.

**Keywords:** Hypothesis test, evidence, Likelihood Ratio, Bayes Factor, invariance.

**PACS:** 02.50.-r

## 1. INTRODUCTION

Simple versus composite hypothesis testing is a general statistical issue in parametric modeling. It consists for a given dataset  $x$  in choosing among the hypotheses

$$H_0 : \theta = \theta_0 ; H_1 : \theta = \theta_1 \quad (1)$$

with  $\theta_0$  pre-assigned and  $\theta_1 \neq \theta_0$  but unknown. Under the classical Bayesian approach adopted here, it means that we use the singular prior  $\delta(\theta - \theta_0)$  on  $H_0$  and we can choose a continuous prior  $\pi(\theta)$ ,  $\theta \in \Theta_1$  on  $H_1$ . This non regular prior, that may appear necessary for hypothesis testing, rises difficulties for classical Bayesian inference. We assume that the data model  $p(x|\theta)$  has the same expression under  $H_0$  and  $H_1$  and does not depend on other parameters than  $\theta$ .

This paper tackles this decision problem using the Posterior distribution of the Likelihood Ratio  $LR = p(x|\theta_0)p(x|\theta)^{-1}$  first advocated as such by [1].

It is organized as follows:

- Section 2 presents standard tests and measures of evidence.
- Section 3 introduces the Posterior distribution of the LR (PLR) and shows it is equal to 1 minus a p-value in a general invariant case.
- Section 4 gives analytical results of the PLR as a distribution (moments, support) and presents it as the posterior probability of a domain.
- Section 5 presents a realistic application of the PLR.

## 2. CLASSICAL TESTS AND MEASURES OF EVIDENCE

### 2.1. Classical Bayesian tests

In addition to the prior  $\pi$  used to define the alternative hypothesis in the test Eq. (1), a prior information  $\Pr(H_0)$  may be used to define the Posterior Odds Ratio by

$$\text{POR}(x) = \frac{\Pr(H_0|x)}{\Pr(H_1|x)} \quad (2)$$

Bayesian hypothesis testing often consists in giving the POR, and thresholding it if a binary decision is required: Reject  $H_0$  if  $\text{POR}(x) \leq \zeta$ . The POR equals the classical Bayes Factor (BF) [5]

$$\text{BF}(x) = \frac{p(x|H_0)}{p(x|H_1)} = \frac{p(x|\theta_0)}{\int p(x|\theta)\pi(\theta)d\theta} \quad (3)$$

up to the multiplicative prior odds ratio  $\Pr(H_0)\Pr(H_1)^{-1}$ . The BF is also classically used for binary decision and its threshold can be interpreted on its own grounds [5].

An important issue of the POR and the BF is that they are not uniquely defined if the prior  $\pi(\theta)$  is improper, ie if  $\int \pi(\theta)d\theta = \infty$ , even though the posterior distribution  $\pi^*(\theta|x)$  is proper. Partial Bayes Factors avoid this difficulty. Among them, the Fractional Bayes Factor  $\text{FBF}(x, b)$ ,  $b \in (0, 1)$  has been proposed in [6] and is defined by:

$$\text{FBF}(x, b) = \frac{p(x|\theta_0)}{\int p(x|\theta)\pi(\theta)d\theta} \left( \frac{p(x|\theta_0)^b}{\int p(x|\theta)^b\pi(\theta)d\theta} \right)^{-1} \quad (4)$$

### 2.2. Frequentist and Bayesian measures of “evidence”

Hypothesis test practitioners in general expect some “predata” information about the operating characteristics of the test. In particular, the significance level of a test, or Probability of False Alarm (PFA), is defined by  $\text{PFA} = \Pr(\text{Reject } H_0|H_0)$ . To calibrate a test by fixing the PFA (e.g. at 5%) is attributed to Neyman-Pearson [7].

Some “postdata” measure of the reliability of a decision is in general also expected. We use the expression “measure of evidence of an hypothesis” like in standard analyses [3, 8]. The most classical frequentist post-data measure about the significance of a decision is the p-value, defined as the lowest PFA for which  $H_0$  would be rejected. For a test of the form “Reject  $H_0$  if  $T(x) \leq \zeta$ ”, the p-value is in general [7] given by:

$$p_{\text{val}}\{T(x)\} = \Pr\{T(y) \leq T(x)|H_0, x\} \quad (5)$$

It can be easily verified that the distribution of the p-value under  $H_0$  is uniform in  $(0, 1)$  if  $T(X)$  is a continuous random variable. In the Bayesian frame, apart from the Bayesian p-value [9],  $\Pr(H_0|x)$  is in general considered as the only Bayesian measure of evidence. In the simple vs composite test, the frequentist p-value and  $\Pr(H_0|x)$  are shown to be irreconcilable in the two-sided test: whatever the prior chosen,  $\Pr(H_0|x) > p_{\text{val}}$  [3].

However, in decision theory,  $\Pr(H_0|x)$  is derived from the estimation problem of the indicator function  $I_{\{\theta_0\}}(\theta)$ , defined by  $I_{\{\theta_0\}}(\theta_0) = 1$  and  $I_{\{\theta_0\}}(\theta) = 0$  if  $\theta \neq \theta_0$ . Under the binary cost function, the Bayes estimate of  $I_{\{\theta_0\}}(\theta)$  consists in thresholding  $\Pr(H_0|x)$  and its Bayes risk is equal to  $\Pr(H_0|x)$  or  $1 - \Pr(H_0|x)$ . Under the quadratic loss function, the Bayes estimate is equal to  $\Pr(H_0|x)$ . Furthermore, [10, 11] claim that for hypothesis testing a measure of evidence should involve the Likelihood Ratio (LR). Consequently, [4] emphasizes that in a formal sense  $\Pr(H_0|x)$  measures the accuracy of a test, not the evidence of an hypothesis.

### 3. PLR AS A MEASURE OF EVIDENCE

#### 3.1. Posterior distribution of the Likelihood Ratio (PLR), review

In the simple vs simple hypothesis test, the Likelihood Ratio (LR) maximizes the Probability of good Detection for a given PFA. Its use as a measure of evidence is straightforward since the parameters under both hypotheses are known.

On the contrary, in the simple vs composite hypothesis test (1),  $\text{LR}(x, \theta)$  is defined by

$$\text{LR}(x, \theta) = \frac{p(x|\theta_0)}{p(x|\theta)} \quad (6)$$

and the parameter  $\theta$  under the alternative hypothesis has to be accounted for. A. Dempster proposed in [1] not to integrate  $\text{LR}(x, \theta)$  over  $\theta$  but to give the probability that  $\text{LR}(x, \theta)$  is lower than some threshold. When the observed dataset  $x$  is the realization of a random process with distribution  $p(x|\theta)$  relying on a fixed parameter  $\theta$  for the experiment performed and under study, instead of comparing  $p(x|\theta_0)$  to  $\int p(x|\theta)\pi(\theta)$ , he advocates to compare “apples to apples” comparing  $p(x|\theta_0)$  to  $p(x|\theta)$ . To account for the other random process, namely  $\theta$  unknown and taking different values on the set of experiments, a cumulative probability can be computed instead of an expectation as done with the BF (see later).

This Posterior cumulative distribution of the Likelihood Ratio under  $H_1$  is noted  $\text{PLR}(x)$ . For a given dataset  $x$ , a binary test relying on  $\text{PLR}(x)$  consists in rejecting

$H_0$  if the posterior probability that the observed data are “much more” likely under  $H_1$  with parameter  $\theta$  than  $H_0$  is “high enough”:

$$\text{Reject } H_0 \text{ if } \text{PLR}(x, \zeta) > p \quad (7)$$

$$\text{with } \text{PLR}(x, \zeta) = \Pr\{\text{LR}(x, \theta) \leq \zeta | x\} \quad (8)$$

The right side of Fig. 1 illustrates the two decision cases:  $H_0$  rejected and  $H_0$  accepted.

Then, the PLR has only been studied as such by M. Aitkin in 1997 [2, 12] and 2005 [13]. According to Dempster and Aitkin, their fundamental results are related to the result found in [1]: the PLR is equal to 1 minus the standard p-value in a normal case with a flat prior when studying the location parameter. Aitkin generalized this result adding a nuisance parameter in the mean, studying asymptotically any density etc, but all generalizations still involved somehow the normal density and its mean.

Note that the PLR is equal to the e-value of the Full Bayesian Significance Test (recently studied by [14]) in the singular case and for a reference prior equal to  $\pi$ . Therefore, the PLR benefits from some properties derived for the e-value and reciprocally some properties derived here can be extended to the e-value.

### 3.2. Invariant case: PLR = 1 - p-val

$\Pr(H_0|x)$  is shown to be irreconcilable with a p-value under general cases [3]. In contrast, the PLR can be shown to be equal to 1 minus a p-value in an invariant setting.

This result generalizes the one of [1, 2]: the likelihood has to belong to an invariant family, the flat prior has to be replaced by the right Haar prior and in the p-value the likelihood is weighted by a modulus.

Invariance is a central framework to unify frequentist, Fisherian pivotal and Bayesian paradigms [15]. See a discussion of the novelty of the result in Sec. 4.2.2. Below are some of the more simple definitions and hypotheses involved in such a result.

**Definition 1** Let  $\mathcal{F}_\Theta = \{f(\cdot|\theta), \theta \in \Theta\}$  a family of densities wrt a measure  $\mu$  on  $\mathcal{X}$  uniquely determined by  $\theta$ . It is said to be invariant under the transformation group  $\mathcal{G}$  if  $Y = g(X)$  has density  $f(\cdot|\theta^*) \in \mathcal{F}_\Theta$  if  $X$  has density  $f(\cdot|\theta) \in \mathcal{F}_\Theta$ . We define  $\bar{\mathcal{G}}$  as the set of all functions  $\bar{g}$  induced by the group  $\mathcal{G}$  with  $\theta^* = \bar{g}(\theta)$ .  $\bar{\mathcal{G}}$  is a group.

In Bayesian models, it is often required to define a “non-informative” prior related to some specific property [16]. In particular, model invariance under a group should be accounted for using the right Haar prior. It is in general improper.

**Definition 2 (Haar measure)** A right invariant Haar measure  $H^r$  on a group  $\mathcal{G}$  is a measure which, for all measurable functions  $\kappa$  with compact support on  $\mathcal{G}$  satisfies

$$\int_{\mathcal{G}} \kappa(g) H^r(dg) = \int_{\mathcal{G}} \kappa(gg_0) H^r(dg) \quad \forall g_0 \in \mathcal{G}$$

A left invariant Haar measure  $H^l$  is defined replacing  $gg_0$  by  $g_0g$ .

We assume that the function  $\phi_\theta : \bar{\mathcal{G}} \rightarrow \Theta$ ,  $\phi_\theta(\bar{g}) = \bar{g}(\theta)$  is isomorphic for all  $\theta$ . The prior measure induced by  $H^r$  is defined for any  $a \in \Theta$  by  $\Pr(\theta \in A) = H^r(\phi_a^{-1}A) \forall A \subset \Theta$ . The measure  $\mu$  on  $\mathcal{X}$  is here simply assumed to be induced by  $H^r$  in the same way.

The statistics involved in the p-value is the likelihood weighted by the modulus of  $\mathcal{G}$ .

**Definition 3** *The modulus of  $\mathcal{G}$  is the function  $\Delta$  defined on  $\mathcal{G}$  to  $(0, \infty)$  which satisfies*

$$\int_{\mathcal{G}} \kappa(gg_0^{-1})H^l(dg) = \Delta(g_0) \int_{\mathcal{G}} \kappa(g)H^l(dg) \quad \forall \kappa, g_0$$

**Theorem 1** *Let  $\mathcal{F}_\Theta = \{f(\cdot|\theta), \theta \in \Theta\}$  be a family of probability densities wrt a measure  $\mu$  on  $\mathcal{X}$ . Assume that:*

1.  $\mathcal{F}_\Theta$  is invariant under the group of transformations  $\mathcal{G}$ .
2.  $\phi_\theta$  and  $\phi_x$  defined above are bijective.  $\mathcal{X}$  and  $\Theta$  are isomorphic.
3. The prior measure on  $\Theta$  is the measure induced by  $H^r$  from  $\phi_\theta$ .
4. The measure  $\mu$  on  $\mathcal{X}$  is the measure induced by  $H^r$  from  $\phi_x$ .

*Then the PLR defined in Eq. (8) can be expressed as the frequentist integral:*

$$PLR(x, \zeta) = \Pr \left\{ \frac{f(x|\theta_0)}{\Delta(\phi_c^{-1}(x))} \leq \zeta \frac{f(y|\theta_0)}{\Delta(\phi_c^{-1}(y))} \mid H_{0,x} \right\} \quad \text{for any } c \in \mathcal{X} \quad (9)$$

The proof of this theorem is available in [17] and its originality is discussed in Sec. 4.2.2. In the Euclidian differentiable case,  $\Delta$  can be replaced by a Jacobian [16].

Setting  $\zeta = 1$  in Eq. (9) leads directly to the following corollary where the second equality follows from the uniform distribution of the p-value under  $H_0$ .

**Corollary 1** *Under the same hypotheses as theorem 1*

$$PLR(x, 1) = 1 - p_{\text{val}} \left\{ \frac{f(x|\theta_0)}{\Delta(\phi_c^{-1}(x))} \right\} \quad (10)$$

*Consequently, the PFA of the test (7) for  $\zeta = 1$  equals  $1 - p$ .*

The assumption of an isomorphism between  $\mathcal{X}$  and  $\Theta$  is restrictive. This constraint is relaxed considering a family of densities where the invariance is defined on  $\tau(X) \in \mathcal{T}$ , a sufficient statistics for  $\theta$ , such that  $\mathcal{T}, \mathcal{G}, \Theta$  and  $\bar{\mathcal{G}}$  are isomorphic. Then, an extension of the proof of theorem 1 leads to (9,10) where the statistics used in the p-value is now  $f_s(\tau|\theta_0)/\Delta(\phi_c^{-1}(\tau))$  where  $f_s(\tau|\theta_0)$  is the marginal distribution of  $\tau(X)$  under  $H_0$ .

## 4. OTHER PROPERTIES AND INTERPRETATIONS OF THE PLR

### 4.1. Moments and support of the PLR: relations to (F)BF and GLR

**Proposition 1** *If the prior  $\pi$  is proper, the posterior moments of the LR equal the FBF:*

$$FBF(x, b) = E[LR(x, \theta)^{1-b} | x] \quad \forall b \in \mathbb{R} \quad (11)$$

In particular,  $\text{BF}(x) = \text{FBF}(x, 0)$  is the posterior mean of the LR. This result, noted for example by Newton and Raftery in 1994, is used and improved for MCMC [18].

The next results assume that the hypotheses are nested:  $\theta_0 \in \text{Sup}(\pi)$ , eventually at the edge. This case is of interest in many practical applications, like in Sec. 5.

Define the Generalized Likelihood Ratio (GLR) as  $\text{LR}(x, \theta)$  evaluated at  $\theta = \hat{\theta}_{\text{ML}}(x)$ :

$$\text{GLR}(x) = \text{LR}(x, \hat{\theta}_{\text{ML}}(x)) = \frac{p(x|\theta_0)}{\max_{\theta} p(x|\theta)} \quad (12)$$

**Proposition 2** *When  $\theta_0 \in \text{Sup}(\pi)$ , the posterior density of the LR verifies:*

- *The minimum of its support is  $\text{GLR}(x)$ :  $\min_{\zeta} \{\zeta : \text{PLR}(x, \zeta) > 0\} = \text{GLR}(x)$*
- *Under regularity assumptions that get stronger as  $L$  (the length of  $\theta$ ) increases, the function  $\zeta \rightarrow \text{PLR}(x, \zeta)$  has an infinite derivative for  $\zeta \rightarrow \text{GLR}(x)^+$ .*

The first property is a direct consequence of Eq. (12). The proof of the second is more delicate in the multivariate case. We show in [17] that if locally there exists  $\alpha > L$  and  $(\alpha_1, \dots, \alpha_L) \in \mathbb{R}_{+*}^L$  such that for all  $\theta$  close enough to  $\hat{\theta}_{\text{ML}}(x)$

$$\text{GLR}(x) < \text{LR}(x, \theta) \leq \text{GLR}(x) + \sum_{\ell=1}^L \alpha_{\ell} (\theta - \hat{\theta}_{\text{ML}}(x))_{\ell}^{\alpha}$$

then  $\varepsilon^{-1} \Pr(\text{GLR}(x) < \text{LR}(x, \theta) \leq \text{GLR}(x) + \varepsilon | x) \rightarrow \infty$  when  $\varepsilon \rightarrow 0$ .

Illustrations of Prop. 1 and Prop. 2 are shown in Fig. 1.

## 4.2. PLR and credible domain

### 4.2.1. PLR test or posterior probability of a domain

Unlike  $\Pr(H_0|x)$ ,  $\text{PLR}(x)$  equals the “observed credible level” of a domain  $C(x) \subset \Theta$ :

$$\begin{aligned} \text{PLR}(x) &= \Pr(\theta \in C(x) | x) \\ \text{with } C(x) &= \{\theta : \text{LR}(x, \theta) \leq \zeta\} \end{aligned} \quad (13)$$

For such a test, the strong post-experimental relationship between tests and credible domains (see [16] for the definition) that was expected in [19] is straightforward. It is similar to the well known pre-experimental equivalence: a domain  $C(x)$  is said to be a  $1 - \alpha$  confidence domain if  $\Pr(\theta \in C(x) | \theta) \geq 1 - \alpha \forall \theta$ . The equivalence  $\theta \in C(x) \Leftrightarrow x \in R(\theta)$  for some rejection region  $R(\theta_0)$  at significance level  $\text{PFA}(\theta_0) \leq \alpha \forall \theta_0$  implies a pre-experimental equivalence of confidence interval and hypothesis test rejection region.

### 4.2.2. Stein theorem relating Bayesian and frequentist inferences

In the group invariant setting involved in Sec. 3.2, if a domain  $C(x) \subset \Theta$  satisfies  $\bar{g}C(x) = C(g(x))$  with  $\bar{g}C(x) = \{\bar{g}\theta, \theta \in C(x)\}$ , then a result from Stein (1965) states

that  $\Pr(\theta \in C(x)|x) = \alpha \forall x$  and  $\Pr(\theta \in C(x)|\theta) = \alpha \forall \theta$ . See for example [16] or [20] for complementary views of this result with simple sets of hypotheses.

However,  $C(x)$  is given by Eq. (13) for the PLR. Then, the domain invariance condition is only satisfied if  $\{\theta_0\}$  is invariant under  $\mathcal{G}$ , i.e.  $\bar{g}\theta_0 = \theta_0 \forall \bar{g} \in \mathcal{G}$ . We study a case incompatible with this assumption:  $\bar{g}\theta_0 = \theta_0$  is impossible if  $\phi_\theta$  used in theorem 1 is bijective. This explains why when the integral over  $\Theta$  switches into an integral over  $\mathcal{X}$  an additional term,  $\Delta(\phi_c^{-1}(x))$ , appears in the domain of integration.

### 4.2.3. Interpretation perspectives

*BF as a point estimate.* The relationship established by the PLR between a test and a credible domain might be of deep interest. The same way a point estimate may not give enough information about a parameter, the same way  $\text{BF}(x)$  seen as the posterior mean of  $\text{LR}(x, \theta)$  may appear as an insufficient inference about it. As an uncertainty it is possible to give the risk of the inference, but for point estimation confidence domains are in general preferred. For testing purposes, the BF is in general thresholded, so that the relevant confidence interval should be constrained to be one-sided. The PLR measures this exact idea (see Fig. 1 for illustration).

*“Post-data hypothesis”.* When  $x$  is the realization of a random process whose distribution is  $p(\cdot|\theta)$ , it might make sense to somehow label  $x$  by  $\theta$ . Reciprocally the parameter  $\theta$  of interest is the one that leads to  $x$ : we make an inference on “ $\theta(x)$ ”. Instead of partitioning  $\Theta$  to define hypotheses,  $\Theta|x$  would then be partitioned. Such a view requires precautions but may lead to no incoherence. In this case, note that if for a given  $x$  the function  $\theta \rightarrow p(x|\theta)$  is bijective,  $\theta = \theta_0$  is equivalent to  $p(x|\theta) = p(x|\theta_0)$ , i.e.  $H_0 : \theta = \theta_0$  (for a given  $x$ ) is equivalent to  $H_0(x) : p(x|\theta) = p(x|\theta_0)$ . However, the fact that the density at  $x$  is equal to the density of  $x$  for the parameter  $\theta_0$  may be not the satisfactory hypothesis to test. Instead we might want to test: “ $\theta$  less likely i.e. more surprising than  $\theta_0$ ”:  $H_0(x) : p(x|\theta) < \zeta p(x|\theta_0)$  In this case,  $\Pr(H_1(x)|x) = \text{PLR}(x, \zeta^{-1})$ .

## 5. PLR IMPLEMENTATION AND REALISTIC APPLICATION

An analytical computation of the PLR is complicated for realistic models, but its numerical computation may be easy using a Monte Carlo Markov Chain (MCMC) algorithm:

1. generate  $\{\theta^{[j]} \sim \pi^*(\theta|x)\}_j$  using a MCMC algorithm
2. compute the chain  $\{\text{LR}(x, \theta^{[j]})\}$
3. compute the PLR (8) as the empirical cumulative distribution of the LR chain
4. if  $H_0$  is rejected, use the chain  $\{\theta^{[j]}\}$  for estimation

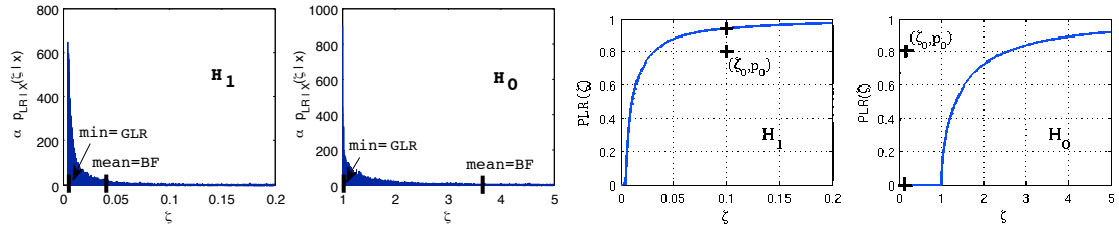
The Markov chain  $\{\text{LR}(x, \theta^{[j]})\}$  can also be used to compute the (F)BF using an importance sampling procedure. Otherwise, see [18] for a better estimate.

This detection procedure is realistically applied to the detection of exoplanets from direct imaging using the future VLT instrument SPHERE. Realistic astrophysical datasets

are simulated by the dedicated Software Package SPHERE [21]. A dataset  $x$  is simulated under  $H_1$  with a luminosity contrast of  $10^6$  between the star and the exoplanet and another under  $H_0$  (no exoplanet).  $x$  consist in  $2 \times 20$  images acquired simultaneously in 2 spectral channels and the  $\theta$  vector refers to the exoplanet intensity in the 2 channels.

A hierarchical Bayesian model dedicated to this context has been developed in [22]. The Markov chain  $\{\theta^{[j]} \sim \pi^*(\theta|x)\}_j$  necessary to compute the test statistics is obtained from a slice sampling method [23]. The procedure is applied to the two datasets.

Fix  $\zeta = 0.1$  and  $p = 0.8$  for the BF and PLR tests. Under the  $H_1$  case, the BF and PLR tests reject  $H_0$ : BF = 0.04 and  $\text{PLR}(x, 0.1) = 0.94$ . In the  $H_0$  case, both tests accept  $H_0$ : BF = 3.7 and  $\text{PLR}(x, 0.1) = 0$ . More information is presented on Fig. 1.



**FIGURE 1.** Left: Histograms of the  $\{\text{LR}(x, \theta^{[j]})\}$  chains. The properties of Sec. 4.1 are displayed. Right: Corresponding cumulative distributions:  $\text{PLR}(x, \zeta)$ . Left:  $H_0$  is rejected; right:  $H_0$  is accepted.

## REFERENCES

1. A. P. Dempster, "The direct use of likelihood for significance testing," in *Proceedings of Conference on Foundational Questions in Statistical Inference*, 1974.
2. M. Aitkin, *Statistics and Computing* (1997).
3. J. Berger, and T. Selke, *Journal of the American Statistical Association* (1987).
4. J. Hwang, G. Casella, C. Robert, M. Wells, and R. Farrell, *Annals of Statistics* (1992).
5. R. Kass, and A. Raftery, *Journal of the American Statistical Association* (1995).
6. A. O'Hagan, *Journal of the Royal Statistical Society* (1995).
7. E. L. Lehmann, and J. P. Romano, *Testing statistical hypotheses*, Springer, 2005, 3rd edn.
8. G. Casella, and R. L. Berger, *Journal of the American Statistical Association* (1987).
9. X.-L. Meng, *Annals of Statistics* (1994).
10. A. Birnbaum, *Journal of the American Statistical Association* (1962).
11. R. Royall, *Statistical evidence: a likelihood paradigm*, Chapman and Hall / CRC Press, 1997.
12. A. P. Dempster, *Statistics and Computing* (1997).
13. M. Aitkin, R. J. Boys, and T. Chadwick, *Statistics and Computing* (2005).
14. W. Borges, and J. Stern, *Logic journal of the IGPL* (2007).
15. M. Eaton, and W. Sudderth, *Bernoulli* (1999).
16. J. O. Berger, *Statistical decision theory and Bayesian analysis*, Springer-Verlag, 1985, 2nd edn.
17. I. Smith, *Détection d'une source faible : modèles et méthodes statistiques. Application à la détection d'exoplanètes par imagerie directe.*, Ph.D. thesis, Université de Nice Sophia-Antipolis (to be defended) (2010).
18. A. E. Raftery, M. A. Newton, J. M. Satagopan, and P. N. Krivitsky, *Bayesian statistics* (2007).
19. C. Goutis, and G. Casella, *Annals of the Institute of Statistical Mathematics* (1997).
20. T. Chang, and C. Villegas, *The Canadian Journal of Statistics* (1986).
21. M. Carillet, A. Boccaletti, et al., "The Software Package SPHERE: a numerical tool for end-to-end simulations of the VLT instrument SPHERE," in *Adaptive Optics Systems*, SPIE, 2008.
22. I. Smith, and A. Ferrari, "Detection from a multi-channel sensor using a hierarchical Bayesian model," in *ICASSP*, 2009.
23. C. P. Robert, and G. Casella, *Monte Carlo Statistical Methods*, Springer-Verlag, 1999.