IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 12, Issue 3 Ver. VII (May. - Jun. 2016), PP 71-78
www.iosrjournals.org

Full and Marginal Bayesian Significance Test: A Frequentist Comparison

Juliet Gratia D'Cunha [1], K. Aruna Rao[2]
[1,2](*Department of Statistics, Mangalore University, India*)

Abstract: *Full Bayesian Significance Test (FBST) was derived by Pereira and Stern for testing a sharp (precise) null hypothesis. The evidence measure of this test has similar interpretation as the p-value of the frequentist significance test. However, it does not suffer from the drawbacks of the p-value. Pereira and his associates argued that full posterior distribution be used to obtain the evidence measure. One of the advantages of the Bayes procedure is that it can handle nuisance parameters easily by integrating this parameter from the posterior distribution. In this paper we derive the Full and the Marginal Bayesian Significance Test (MBST) for testing the specified value of the median for the two parameter lognormal distribution. Simulation is used to compare the FBST and the MBST in terms of power of the test. The results indicate that the critical value of the evidence measure for the MBST is closely related to the p-value of the t test. In terms of evidence, FBST provides stronger evidence against the null hypothesis compared to the MBST and the t test. Although t test is more powerful compared to the FBST and MBST, there is no difference in the power of FBST and MBST. An illustrative example is also provided.*

Keywords: *Full Bayesian Significance Test, lognormal distribution, power of the test, p-value*

I. Introduction

Full Bayesian Significance Test (FBST) was proposed by Pereira and Stern [1]. It was done with the following objectives:

a) To propose a Bayesian Significance Test for testing a sharp (or precise) null hypothesis.
b) To propose a Bayesian evidence measure (Ev) against the null hypothesis which is similar to the frequentist p-value.

In a series of paper, Pereira and his associates studied the various properties of FBST and the advantage of the evidence measure over the p-value. These properties include the invariance property of the FBST under coordinate transformation and that FBST is a Bayes rule under an appropriate loss function (see [2]). With an appropriate reasoning Pereira et al. [3] argue that FBST is a significant test for testing a sharp null hypothesis. Other papers on FBST are the following: Pereira and Stern [4] use FBST for selecting a model in a class of nested models, they suggest that FBST can be easily implemented using modern numerical optimization and integration techniques and also indicate that FBST needs no additional assumptions. Full Bayesian Significance Test for coefficient of variation (CV) was considered by Pereira and Stern [5]. They use FBST to compare CV in applications arising in finance and industrial engineering. FBST for testing the hypothesis of independence in a Holgate Bivariate Poisson distribution was developed by Stern and Zacks [6]. They compare the power of the test with several well known classical tests (Non-Bayesian) and using Monte Carlo simulation showed that FBST has better power properties compared to the other tests.

While proposing FBST, Pereira and Stern [1] strongly advocated the use of full posterior distribution for computing the evidence measure. The word 'full' refers to the joint posterior distribution of all the parameters in the model. In their paper Pereira et al. [3] argue that the marginal approach should not be used for computing the Ev measure. They cite Basu [7] and Good [8] to support their claim on the Bayesian inference based on full posterior distribution. FBST shares some similarity with likelihood ratio test (LRT) (see section 2 for details) and is appealing even for a frequentist. However, the insistence that Ev measure be computed based on the full posterior distribution is little disturbing for frequentist as well as some section of the Bayesians.

Inference based on the marginal posterior distribution for the parameters of interest is well accepted by many Bayesians. This is the standard method of eliminating nuisance parameters in Bayesian inference and Bayesians cite it as an advantage of the procedure compared to the ad hoc procedure in frequentist approach (see [9], [10] and [11]). Dawid [12] and Severini [13] and the references cited therein, use marginal posterior approach to justify "the ad hoc procedure of the frequentist". Although Bayesians call this as ad hoc procedure, the conditioning and invariance approach have played a vital role in classical inference. When Bayesians also approve the marginal approach, a natural question by the frequentist is that "What is the superiority of the FBST compared to the MBST (Marginal Bayesian Significance Test)?" In MBST the marginal posterior distribution of the parameter of interest is used to compute the evidence measure. Pereira and Stern [1] have indicated that the evidence measure is stronger for testing a real parameter compared to the case of testing the same in the presence of nuisance parameter. This has motivated us to make a frequentist comparison of the FBST and the MBST. In this comparison we use the theoretical argument as well as the evidence from a simulation study. The results indicate that, in general there is no computational advantage for the MBST, while from the Bayesian perspective the evidence in FBST is stronger compared to the evidence measure in MBST.

The paper is organized as follows: Section 2 describes FBST and MBST. In Section 3, the problem of testing for the median of the two parameter lognormal distribution is discussed. Numerical power comparison of the FBST and MBST is presented in Section 4. Section 5 presents results and discussions while Section 6 takes up the analysis of a real life data. The paper concludes in Section 7.

II. Full And Marginal Bayesian Significance Test

2.1 Full Bayesian Significance Test (FBST)

Let X be a random variable with density $f(x|\theta), \theta$ real. Let $x = (x_1, \dots, x_n)$ denote the realized value of the sample. Let $\pi(.)$ be a generic notation of a probability density. The problem of interest is to test $H_0\colon \theta = \theta_0$. Let $\pi(\theta)$ denote the prior density of θ and $\pi(\theta|x)$ denote the posterior density given the data. Let $\gamma = P(\theta\colon \pi(\theta|x) > \pi(\theta_0|x))$. The evidence measure Ev is defined as $Ev = 1 - \gamma$. In FBST the null hypothesis is rejected for small values of Ev.

To understand the similarity between FBST and the likelihood ratio test (LRT), in fig 2.1a and 2.1b, we have plotted separately the posterior density $\pi(\theta|x)$ and the density of the Likelihood ratio statistic $T_0(X) = \frac{L(\theta_0;x)}{L(\hat{\theta};x)}$, where $\hat{\theta}$ is the maximum likelihood estimator (MLE) of θ.

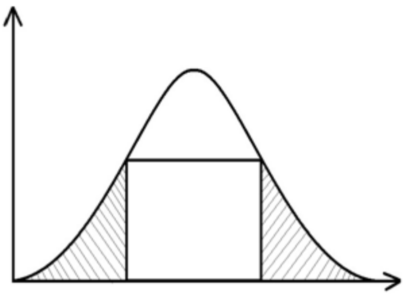

Fig 2.1 a. Posterior Density $\pi(\theta|x)$

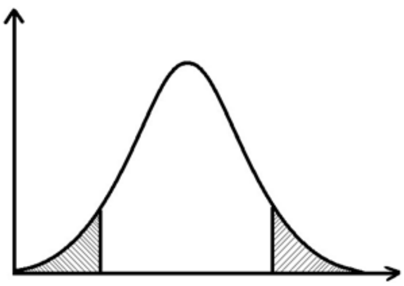

Fig 2.1 b. Density of Likelihood Ratio Statistic

In fig 2.1a, the unshaded region corresponds to the tangential set for the null hypothesis and γ corresponds to the probability of this tangential set. In fig 2.1b, the shaded region is the region of rejecting the null hypothesis. The probability of the tail areas of the posterior distribution of θ is the Ev measure while the probability of the tail areas of the density function of $T_0(.)$ refers to the p-value. The figures clearly prove the closeness between the logical reasoning of FBST and the LRT. This idea is further elaborated in Section 3.

Let θ be vector valued, $\theta = (\psi, \lambda)$ where ψ be real and λ may be real or vector valued. The hypothesis of interest $H_0\colon \psi = \psi_0$and γ is now defined as $\gamma = P(\psi, \lambda\colon \pi(\psi, \lambda) > sup_\lambda \pi(\psi_0, \lambda))$. The Ev measure, as in previous case is $Ev = 1 - \gamma$. In multi parameter set up the plotting of posterior density and the likelihood function is difficult. Nevertheless, if one looks at the expression for γ, one can realize the closeness between the FBST and the LRT.

2.2 Marginal Bayesian Significant Test (MBST)

Pereira and his associates did not propose the Bayesian Significance Test based on marginal posterior distribution. Using the previous notation, the null hypothesis of interest is $H_0\colon \psi = \psi_0$. We can define the evidence measure based on the marginal posterior distribution $\pi(\psi|x)$ using the same logic of Pereira and Stern [1]. Let $\gamma_1 = P\big(\psi\colon \pi(\psi|x) > \pi(\psi_0|x)\big)$ and the Ev measure is given by $Ev_1 = 1 - \gamma_1$.

The computation of FBST involves two steps: a) Optimization and b) Integration. It is computationally very tedious, while the MBST involves only one step, i.e integration. One can use MCMC for computation of γ (FBST) and γ_1(MBST). The generation of samples from a multivariate posterior distribution is computationally more tedious than generating observations from a univariate posterior distribution. In order to overcome the computational burdens Cabras et al. [14] obtained higher order approximation for the evidence measure in MBST and thereby for FBST. In Section 3 we show the closeness between the MBST and UMPU test for testing parameters of the lognormal (in turn normal) distribution.

III. FBST and MBST for Testing the Specified Value of the Median in Lognormal Distribution

3.1 Notations and test procedure

Given a random sample of size n from two parameter lognormal distribution with log location parameter μ and log scale parameter σ, $\bar{Z}$ and ${S_z}^2$ are sufficient statistic for μ and σ^2 where $Z_i = \log x_i, i = 1, \dots n$ and $\bar{Z}$ and ${S_z}^2$ denote the sample mean and sample variance, respectively for the transformed variable Z. The likelihood for μ and σ^2 is given by,

$$L(\mu, \sigma^2; z) = \frac{1}{\sqrt{2\pi}\frac{\sigma}{\sqrt{n}}} e^{-\frac{1}{2}\frac{(\mu-\bar{z})^2}{\sigma^2/n}} \frac{\left(\frac{(n-1){S_z}^2}{2}\right)^{\frac{n+1}{2}}}{\Gamma\left(\frac{n+1}{2}\right)} \left(\frac{1}{\sigma^2}\right)^{\left(\frac{n+1}{2}\right)-1} e^{-\frac{1}{2}\left(\frac{n-1}{\sigma^2}\right){S_z}^2}, -\infty < \mu < \infty, \sigma^2 > 0 \qquad (1)$$

Lognormal distribution was the topic of interest in the papers by Zellner [15], Zhou and Tu [16], Krishnamoorthy and Mathew [17], Harvey et al. [18], Harvey and Merwe ([19],[20]) and D'Cunha and Rao ([21],[22]). The commonly used priors are: i) Right invariant prior $\pi_1(\mu, \sigma) = \frac{1}{\sigma}$ ii) Left invariant Jeffreys prior $\pi_2(\mu, \sigma) = \frac{1}{\sigma^2}$ iii) Jeffreys rule prior $\pi_3(\mu, \sigma) = \frac{1}{\sigma^3}$ and iv) Uniform prior $\pi_4(\mu, \sigma) = 1$. Jeffreys rule prior is proportional to $|I(\mu, \sigma^2)|^{1/2}$, where $I(\mu, \sigma^2)$ is the Fisher information matrix for the normal distribution. This prior was used in the past by Harvey and Merwe [20]. Some of these priors were used in the past by Zellner [15], Harvey et al. [18], Harvey and Merwe ([19],[20]) and D'Cunha and Rao ([21], [22]).The additional prior used in the present investigation is the probability matching prior $\pi_5(\mu, \sigma) = \sigma^2$. Although this prior is not commonly used, the use of this prior leads to agreement between the UMPU/UMP invariant test and the Marginal Bayesian Test based on the *t* statistic defined on the parameter space. The posterior distribution is the product of gamma for $\eta = \frac{1}{\sigma^2}$ and the conditional normal distribution for μ. To save space we give below the posterior distribution of μ and σ^2 for the right invariant prior.

$$\pi_1(\mu, \sigma|z) = \frac{1}{\sqrt{2\pi}\frac{\sigma}{\sqrt{n}}} e^{-\frac{1}{2}\frac{(\mu-\bar{z})^2}{\sigma^2/n}} \frac{\left(\frac{(n-1){S_z}^2}{2}\right)^{\frac{n+2}{2}}}{\Gamma\left(\frac{n+2}{2}\right)} \left(\frac{1}{\sigma^2}\right)^{\left(\frac{n+2}{2}\right)-1} e^{-\frac{1}{2}\left(\frac{n-1}{\sigma^2}\right){S_z}^2}, -\infty < \mu < \infty, \sigma^2 > 0, \qquad (2)$$

(using right invariant prior)

The difference with respect to various priors is only in the shape parameter of the gamma distribution for $\eta = \frac{1}{\sigma^2}$ in the posterior distribution and is symbolically given below,

$\pi_2(\mu, \sigma|z) = N(\mu; \bar{Z}, \sigma^2/n) \times G_\eta\left(\frac{n+3}{2}, \left(\frac{n-1}{2}\right){S_z}^2\right)$ (using left invariant prior) (3)

$\pi_3(\mu, \sigma|z) = N(\mu; \bar{Z}, \sigma^2/n) \times G_\eta\left(\frac{n+4}{2}, \left(\frac{n-1}{2}\right){S_z}^2\right)$ (using Jeffreys Rule prior (4)

$\pi_4(\mu, \sigma|z) = N(\mu; \bar{Z}, \sigma^2/n) \times G_\eta\left(\frac{n+1}{2}, \left(\frac{n-1}{2}\right){S_z}^2\right)$ (using Uniform prior) (5)

$\pi_5(\mu, \sigma|z) = N(\mu; \bar{Z}, \sigma^2/n) \times G_\eta\left(\frac{n-1}{2}, \left(\frac{n-1}{2}\right){S_z}^2\right)$ (using Probability matching prior) (6)

where $z = z_1, z_2, \dots, z_n$. The median of the lognormal distribution is e^μ and testing for the specified value of the median is equivalent to testing $H_0: \mu = \mu_0$. For this hypothesis UMPU/UMP invariant test exists (see [23]). And the test statistic is given by $t = \frac{\bar{Z}-\mu}{S_z/\sqrt{n}}$. Under H_0, t follows central t distribution with $(n-1)$ degrees of freedom (d.f). The null hypothesis is rejected when $|t| > t_{\alpha/2}(n-1)$ where $t_{\alpha/2}(n-1)$refers to the upper $\frac{\alpha}{2}th$ percentile value of t distribution with $(n-1)$ degrees of freedom. For any posterior distribution $\pi_i(\mu, \sigma^2)$ the evidence measure against the null hypothesis is $Ev = 1 - \gamma$ where γ is given by $\gamma = P\left(\mu, \sigma^2 : \pi_i(\mu, \sigma^2) > sup_{\sigma^2}\pi_i(\mu_0, \sigma^2)\right)$. Closed form solutions exist for σ^2 which maximizes the posterior density of $\pi_i(\mu_0, \sigma^2)$. For the frequentist comparison, the test statistic is the evidence measure $Ev = 1 - \gamma$, for which no closed form solution exists. In Section 4, Importance Sampling Approach is used to generate observations from the posterior distribution and to carry out the Monte Carlo Integration.

The marginal posterior density function of $\pi_5(\mu|z)$ for the probability matching prior is given below,

$$\pi_5(\mu|data) = \frac{\sqrt{n}}{\sqrt{2\pi}} \frac{\left\{\left(\frac{n-1}{2}\right){S_z}^2\right\}^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} \frac{\Gamma\left(\frac{n}{2}\right)}{\left\{\frac{1}{2}\left(n(\mu-\bar{Z})^2 + (n-1){S_z}^2\right)\right\}^{n/2}} \qquad (7)$$

We observe that $t_1 = \frac{\mu-\bar{Z}}{S_z/\sqrt{n}}$ follows t distribution with $(n-1)$ d.f. The distinction between the distribution of t and t_1 is that, t is defined on the sample space while t_1 is defined on the parameter space. To save space the marginal posterior distribution of μ for the other priors is not shown here. Let $t' = \frac{\mu-\bar{Z}}{S_z/\sqrt{n}}\frac{\sqrt{v}}{\sqrt{n-1}}$, the marginal posterior distributions of t' are central t distribution with degrees of freedom ν equal to $n+2$ (for Right invariant), $n+3$ (for Left invariant Jeffreys), $n+4$ (for Jeffreys rule), $n+1$ (for Uniform) priors, respectively. Thus the marginal posterior distribution of μ is the location shifted scaled t distribution for all the priors. The evidence measure for the MBST for any prior is given by $P\left(\mu: \pi_i(\mu|z) < \pi_i(\mu_0|z)\right)$, which is also the test statistic for frequentist comparison. Although this probability can be evaluated analytically, we have used Monte Carlo integration so as to have the uniform accuracy for the FBST and the MBST in Section 4.

IV. Finite Sample Comparison on the Performance of FBST, MBST and *t* test

In this Section we compare the power of the FBST, MBST and t test for finite (small) samples. For this purpose an extensive simulation is carried out.

4.1 Simulation experiment for the estimation of the critical values

For a given sample we have to estimate the evidence measure using Monte Carlo Integration by generating observations from the posterior distribution. This constitutes the inner simulation and corresponds to one estimate of the evidence measure. For estimating the critical values, the experiment has to be repeated by generating observations from the normal distribution with the same values of the parameters. This constitutes the outer simulation.

A sample of size n is generated from normal distribution with $\mu = \log(1000)$ and a given value of σ^2. For this sample, the observations for μ and σ^2 is generated from the posterior distribution $\pi(\mu, \sigma^2|data)$. Since the posterior distribution is the product of gamma distribution for $\eta = \frac{1}{\sigma^2}$ and the conditional normal distribution given η, observation for η is generated from the gamma distribution; and using this value of η, the observation for μ is generated from the conditional normal distribution. This constitutes a pair of observation (μ, η). Thus for each sample we are generating 10,000 observations from the posterior distribution. Using the generated values of (μ, σ^2) we have estimated the number of times $\pi(\mu, \sigma^2) > sup_{\sigma^2}\pi(\mu_0, \sigma^2)$.This constitutes the estimated value of γ and $Ev = 1-\gamma$. Using 1000 samples, the Ev measure is obtained. The lower α^{th} percentile of the simulated distribution of the Ev measure is noted down. This is the estimated critical value for the evidence measure in FBST.

For the MBST, for generating observations from the marginal posterior distribution, observations are generated from the $t$distribution with ν degrees of freedom. The generated observation for μ is obtained by the relation, $t = \frac{(\mu-\bar{z})}{s_z/\sqrt{n}}\frac{\sqrt{\nu}}{\sqrt{n-1}}$, where t denotes the generated value from the t distribution and ν denotes the degrees of freedom and $\nu = n+2$ (Right Invariant), $\nu = n+3$ (Left Invariant), $\nu = n+4$ (Jeffreys Rule), $\nu = n+1$ (Uniform) and $\nu = n-1$ (Probability Matching) priors, respectively. Using these 10,000 values of μ, $\gamma_1$is computed by estimating the proportion of times $\pi(\mu|data) > \pi(\mu_0|data)$. By generating 1000 samples from normal distribution the critical value of the Ev_1 is estimated.

The sample sizes considered are $n = 10, 20, 40, 60, 80, 100, 150$ and 200. Various values of σ^2 are used in the simulation. Since σ^2 is related to the coefficient of variation (CV) of the lognormal distribution by the relation $CV = \left(e^{\sigma^2}-1\right)^{1/2}$, the values for σ^2 is estimated corresponding to the CV values of 0.1, 0.3, 0.5, 0.7, 1, 1.5, 2 and 2.5. The level of significance is fixed at $\alpha = 0.05$.

4.2 Estimation of the power of the test

For the estimation of the power of the test, observations from the normal distribution are generated for various values of μ by fixing the value of σ^2 as in the null hypothesis. The power of the test corresponds to the proportion of times the Ev value under the alternative hypothesis is less than or equal to the estimated critical value. The estimation of $Ev$for the alternative hypothesis is similar to the null hypothesis. The power of the *t* test is computed using the critical values of the *t* distribution as this test is exact.

V. Results And Discussion

5.1 Estimated critical values for the FBST and the MBST

Table 5.1 and 5.2 presents the estimated critical values for the FBST and MBST, when $\alpha = 0.05$, for the sample sizes *n* = 10, 20, 60 and 100 and for various combinations of CV. To save space the critical values for other sample sizes are not reported.

Table 5.1: Estimated critical values for FBST for various combinations of CV and sample sizes when α=0.05

Sample Size	Prior	Critical Values for Evidence measure when CV equal to							
		0.1	**0.3**	**0.5**	**0.7**	**1**	**1.5**	**2**	**2.5**
10	**Right Invariant**	0.0730	0.0612	0.0730	0.0612	0.0683	0.0577	0.0575	0.0826
	Left Invariant	0.0559	0.0494	0.0559	0.0494	0.0515	0.0446	0.0462	0.0661
	Jeffreys Rule	0.0410	0.0385	0.0410	0.0385	0.0398	0.0340	0.0333	0.0507
	Uniform	0.1003	0.0902	0.1003	0.0902	0.0952	0.0871	0.0863	0.1166
	Probability matching	0.1412	0.1267	0.1446	0.1256	0.1297	0.1521	0.1182	0.1468
20	**Right Invariant**	0.1079	0.0938	0.1079	0.1079	0.0938	0.1344	0.0743	0.1017
	Left Invariant	0.0979	0.0817	0.0979	0.0979	0.0817	0.1229	0.0719	0.0909
	Jeffreys Rule	0.0852	0.0754	0.0852	0.0852	0.0754	0.1071	0.0580	0.0820
	Uniform	0.1209	0.1085	0.1209	0.1209	0.1085	0.1472	0.0908	0.1209
	Probability matching	0.1293	0.1336	0.1363	0.1293	0.1363	0.1316	0.1403	0.1514
60	**Right Invariant**	0.1348	0.1237	0.1741	0.1231	0.1417	0.1477	0.1650	0.1024
	Left Invariant	0.1275	0.1248	0.1695	0.1220	0.1350	0.1415	0.1602	0.0988
	Jeffreys Rule	0.1261	0.1187	0.1640	0.1153	0.1321	0.1399	0.1585	0.0963
	Uniform	0.1411	0.1321	0.1788	0.1300	0.1471	0.1545	0.1689	0.1088
	Probability matching	0.1374	0.1402	0.1330	0.1876	0.1404	0.1514	0.1382	0.0920
100	**Right Invariant**	0.1320	0.1239	0.1452	0.1320	0.1239	0.1239	0.1718	0.1452
	Left Invariant	0.1278	0.1218	0.1399	0.1278	0.1218	0.1218	0.1705	0.1399
	Jeffreys Rule	0.1179	0.1181	0.1433	0.1179	0.1181	0.1181	0.1647	0.1433
	Uniform	0.1350	0.1273	0.1489	0.1350	0.1273	0.1273	0.1757	0.1489
	Probability matching	0.1471	0.1443	0.1400	0.1441	0.1483	0.1600	0.1081	0.1285

From table 5.1 we notice that for a small sample size n=10 and all the priors the critical values for the FBST ranges from 0.03 to 0.15 while it is 0.05 to 0.15 for n=20. As the sample size increases the critical values ranges from 0.10 to 0.18 for sample size n =60 and it ranges from 0.12 to 0.18 for sample size n =100. No systematic pattern can be identified in the critical values as the CV increases from 0.1 to 2.5. For each prior, as the sample size increases, the estimated critical values converge to a point which is slightly higher than 0.10. This conclusion is based on all the sample sizes and not necessarily to the values reported in the table.

Table 5.2: Estimated critical values for MBST for various combinations of CV and sample sizes when α=0.05

Sample size	Prior	Critical Values for Evidence measure when CV equal to							
		0.1	**0.3**	**0.5**	**0.7**	**1**	**1.5**	**2**	**2.5**
10	**Right Invariant**	0.0163	0.0199	0.0182	0.0173	0.0174	0.0173	0.0206	0.0180
	Left Invariant	0.0132	0.0166	0.0139	0.0123	0.0130	0.0139	0.0155	0.0132
	Jeffreys Rule	0.0100	0.0118	0.0101	0.0094	0.0103	0.0111	0.0121	0.0113
	Uniform	0.0224	0.0276	0.0259	0.0221	0.0238	0.0223	0.0267	0.0244
	Probability matching	0.0392	0.0470	0.0423	0.0390	0.0423	0.0415	0.0469	0.0427
20	**Right Invariant**	0.0379	0.0319	0.0303	0.0292	0.0256	0.0316	0.0393	0.0292
	Left Invariant	0.0327	0.0276	0.0271	0.0286	0.0223	0.0290	0.0340	0.0252
	Jeffreys Rule	0.0286	0.0246	0.0242	0.0239	0.0191	0.0254	0.0306	0.0223
	Uniform	0.0438	0.0364	0.0338	0.0353	0.0291	0.0356	0.0445	0.0328
	Probability matching	0.0531	0.0469	0.0453	0.0464	0.0381	0.0491	0.0536	0.0430
60	**Right Invariant**	0.0443	0.0514	0.0378	0.0455	0.0548	0.0548	0.0405	0.0342
	Left Invariant	0.0420	0.0490	0.0372	0.0453	0.0559	0.0559	0.0401	0.0329
	Jeffreys Rule	0.0401	0.0475	0.0349	0.0399	0.0500	0.0500	0.0374	0.0303
	Uniform	0.0438	0.0525	0.0389	0.0486	0.0574	0.0574	0.0431	0.0357
	Probability matching	0.0493	0.0560	0.0432	0.0531	0.0622	0.0622	0.0482	0.0413
100	**Right Invariant**	0.0422	0.0410	0.0404	0.0400	0.0463	0.0427	0.0425	0.0466
	Left Invariant	0.0426	0.0402	0.0387	0.0392	0.0450	0.0426	0.0409	0.0462
	Jeffreys Rule	0.0411	0.0408	0.0377	0.0393	0.0451	0.0404	0.0406	0.0461
	Uniform	0.0445	0.0422	0.0420	0.0396	0.0473	0.0421	0.0426	0.0483
	Probability matching	0.0451	0.0446	0.0442	0.0442	0.0501	0.0436	0.0455	0.0504

From table 5.2 we observe that for all the priors the critical values for MBST for a small sample size n=10 ranges from 0.01 to 0.05 while it is 0.02 to 0.05 for n =20. As sample size increases the critical value ranges from 0.03 to 0.06 and 0.04 to 0.05 for sample sizes n =60 and 100, respectively. For each prior, as sample size increases, the estimated critical values converge to a point which is slightly higher than 0.05.

A significant conclusion that emerges from the tables is that the critical values for the FBST as well as MBST is higher for the probability matching prior compared to other priors. This is followed by the critical values for Uniform, Right invariant, Left invariant Jeffreys and Jeffreys Rule prior, respectively. It also follows that the evidence provided by the MBST is close to the p-value while the evidence provided by FBST is much stronger compared to the MBST and the p-value.

5.2 Power comparison

The power curve for the FBST and MBST for the 5 priors along with *t* test when the sample size is *n*=100 and CV=0.1 is presented in figures 5.1 to 5.5. The power curves are also estimated when CV=1 and 1.5 and to save space it is not shown here.

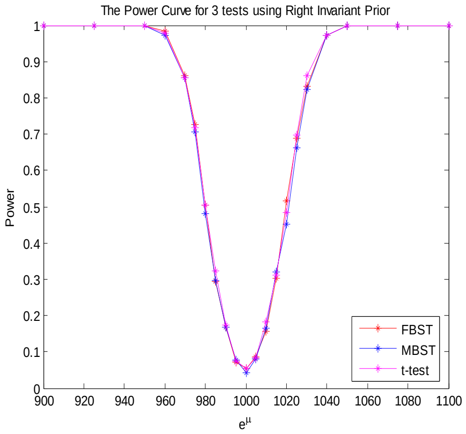

Fig 5.1: Power Curve for Right Invariant prior when *n*=100, CV=0.1

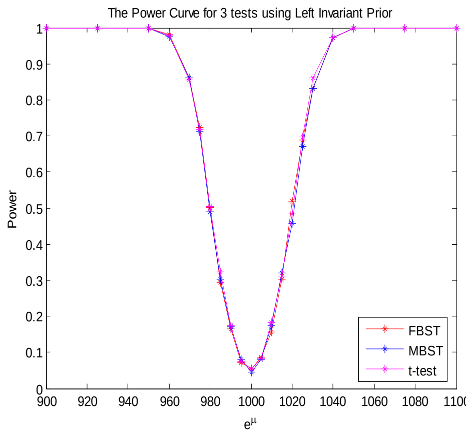

Fig 5.2: Power Curve for Left Invariant prior when *n*=100, CV=0.1

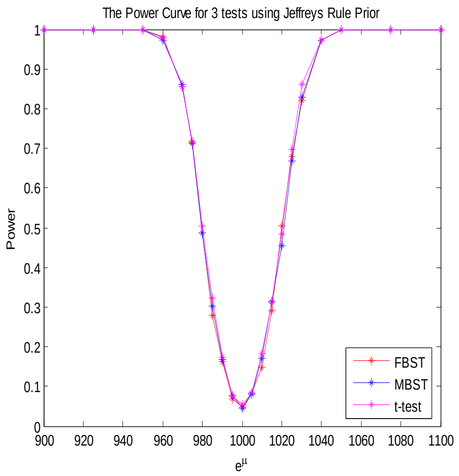

Fig 5.3: Power Curve for Jeffreys Rule prior when *n*=100, CV=0.1

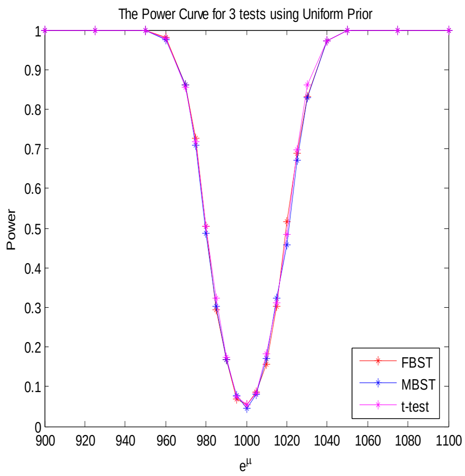

Fig 5.2: Power Curve for Uniform prior when *n*=100, CV=0.1

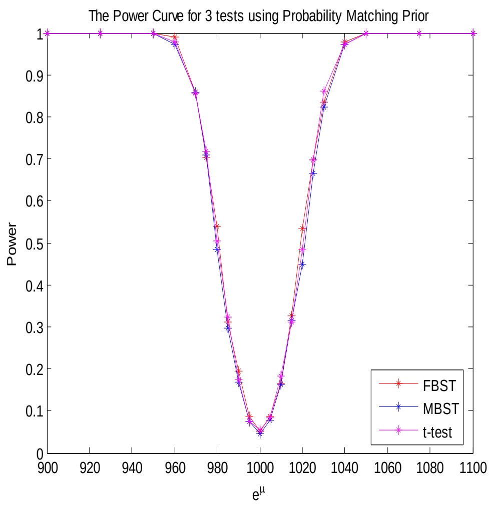

Fig 5.5: Power Curve for Probability Matching prior when *n*=100, CV=0.1

From the figures it follows that for the left and right contiguous alternatives the power of the *t* test is marginally higher compared to the FBST and MBST. For these alternatives the power curves for the FBST and MBST fluctuates from each other marginally. This indicates that among FBST and MBST no test is more powerful than the other. The rate of convergence of the power curves to the value 1 is equal for all the 3 tests.

VI. Example

(Source:http://www.nseindia.com/products/content/equities/equities/eq_security.htm)

To illustrate the use of the tests mentioned in this paper, we have considered a real data set on stock prices. The script chosen for the analysis is TATA STEEL which is listed in the National Stock Exchange (NSE) Limited, India. The data is considered for the period May and June, 2015. The average price per day is taken as the price of the stock. The data is available from the above cited link. The analysis is carried out to check whether the median stock price of the script TATA STEEL for the month of June 2015 is same as that for the month of May 2015.

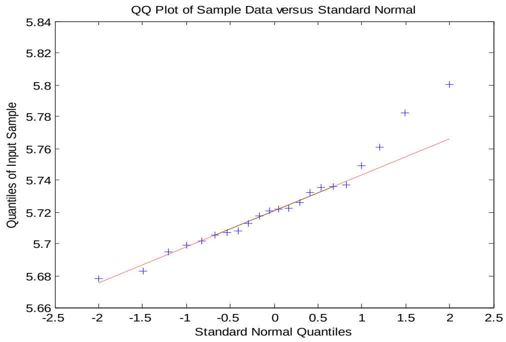

Fig 6.1: Q-Q plot for log transformed stock price data for the script TATA STEEL

The Q-Q plot for the log transformed data is presented in fig 6.1, which confirms that the underlying distribution is normal and there by implies that the underlying distribution for stock prices is lognormal. The median value for the month of May 2015 is Rs. 361.035 and the null hypothesis of interest is that,

$$H_0\text{:}Median\ stock\ price\ for\ the\ month\ of\ June\ 2015,$$
$$i.e\ e^{\mu} = 361.035\ or\ equivalently\ \mu = 5.8890$$

Table 6.1 presents the evidence measure and critical values for the FBST and MBST along with the calculated p-value for the *t* test.

Table 6.1: Computed and critical values of the test statistic

Test	Prior	Value of the test Statistic	Critical value	Decision
FBST	**Right invariant**	0	0.1017	Reject
	Left invariant	0	0.0909	Reject
	Jeffreys rule	0	0.0820	Reject
	Uniform	0	0.1209	Reject
	Probability matching	0	0.1514	Reject
MBST	**Right invariant**	0	0.0292	Reject
	Left invariant	0	0.0252	Reject
	Jeffreys rule	0	0.0223	Reject
	Uniform	0	0.0328	Reject
	Probability matching	0	0.0430	Reject
T-test	-	-26.6785	$t_{0.025}(21)$ =2.0796	Reject

* Critical value refers to sample size *n*=20 and CV=2.5 and the p-value for the *t* test is zero.

Thus the conclusion that arises from the three tests is to reject the null hypothesis. The p-value for the *t* test as well as the evidence measure for FBST and MBST are all equal to zero. This indicates that we cannot distinguish the three tests in terms of evidence against the null hypothesis.

VII. Conclusion

The purpose of this paper is to compare the performance of the Full Bayesian Significance Test (FBST) and the Marginal Bayesian Significance Test (MBST) for testing the specified value of the median for the two parameter lognormal distribution. From a frequentist point of view, the MBST is preferable compared to the FBST as it is related to the UMPU/UMP invariant test as in the present context. Nevertheless, extensive simulation carried out in the paper indicates that there is no difference in the power of the FBST and the MBST for contiguous alternatives for all the 5 priors used in the study. However, from the Bayesian perspective the evidence provided by the FBST is stronger compared to the MBST.

Although there may be valid reasons for using the full model compared to the marginal posterior distribution, we are of the opinion that no justification for this can be given from the frequentist view point. MBST enjoys computational advantage over FBST when closed form solution exists for the marginal posterior distribution. For the problem under consideration as well as from a general scenario, we notice that it is simpler to obtain the third order evidence measure for the MBST compared to the FBST (see [14] for details). In the present context the p-value of the *t* test and thereby the Likelihood Ratio Test (LRT) is closely related to the *Ev* value of the MBST. Although we do not have a proof, we conjecture that the closeness is true in general. Frequentist comparison of the FBST and MBST in non-standard situations is a problem for future research. The program for the simulation is written in Matlab software version 7.0.

Acknowledgements

The first author would like to take the opportunity to thank Government of India, Ministry of Science and Technology, Department of Science and Technology, New Delhi, for sponsoring her with the award of INSPIRE fellowship, which enabled her to carry out this research programme. She is much honoured to be the recipient of this award.

References

[1] C. A. B. Pereira, and J. M. Stern, Evidence and Credibility: Full Bayesian Significance Test for Precise Hypotheses, *Entropy*, *1*, 1999, 99-110.

[2] M. R. Madruga, C. A. B. Pereira, and J. M. Stern, Bayesian evidence test for precise Hypotheses, *Journal of Statistical Planning and Inference*, *117*, 2003, 185-198.

[3] C. A. B. Pereira, J. M. Stern, and S. Wechsler, Can a significance test be genuinely Bayesian? *Bayesian Analysis*, *3 (1)*, 2008, 79-100.

[4] C. A. B. Pereira, and J. M. Stern, Full Bayesian significance test for Coefficient of Variation, *In Bayesian Methods with Applications to Science, Policy, and Official Statistics*, EI George (editor), Creta, Greece: ISBA, 2001, 391-400.

[5] C. A. B. Pereira, and J. M. Stern, Model Selection: Full Bayesian approach, *Environmetrics*, *12(6)*, 2001, 559-568.

[6] J. M. Stern, and S. Zacks, Testing the independence of Poisson variate under the Holgate Bivariate distribution: the power of a new evidence test, *Statistics and Probability Letters*, *60*, 2002, 313-320.

[7] D. Basu, On the elimination of nuisance parameters, *J. Amer. Statist. Assoc.*, *72*, 1977, 355-366.

[8] I. Good, *Good Thinking: the Foundations of Probability and its Applications* (MN: University of Minnesota Press, Minneapolis, 1983).

[9] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis (Second Edition)* (Springer-Verlag, New York, 1985).

[10] G. E. P. Box, and G. C. Tiao, *Bayesian inference in statistical analysis* (John Wiley & Sons, New York, 1992).

[11] J. K. Ghosh, M. Delampady, and T. Samanta, *An introduction to Bayesian Analysis: Theory and methods* (Springer texts in Statistics, New York, 2006).

[12] A. P. Dawid, A Bayesian look at nuisance parameters, *Trabajos de Estadsitica Y de Investigacion Operativa*, *31(1)*, 1980, 167-203.

[13] T. A. Severini, On the relationship between Bayesian and non Bayesian elimination of nuisance parameters, *Statistica Sinica*, *9*, 1999, 713-724.

[14] S. Cabras, W. Racugno, and L. Ventura, Higher order asymptotic computation of Bayesian significance tests for precise null hypotheses in the presence of nuisance parameters, *Journal of Statistical Computation and Simulation*, *85(15)*, 2015, 2989-3001.

[15] Zellner, Bayesian and Non-Bayesian Analysis of the Lognormal Distribution and Lognormal Regression, *Journal of the American Statistical Association, 66(334)*, 1971, 327-330.

[16] X. H. Zhou, and W. Tu, Interval estimation for the ratio in means of log-normally distributed medical costs with zero values, *Computational Statistics and Data Analysis*, *35*, 2000, 201–210.

[17] Krishnamoorthy, K., and Mathew, T. Inferences on the means of lognormal distributions using generalized P-values and generalized confidence intervals, *Journal of Statistical Planning and Inference*, *115*, 2003, 103 – 121.

[18] J. Harvey, P. C. N. Groenewald, and A. J. V. Merwe, Bayesian confidence intervals for the mean of a lognormal distribution: A comparison with the MOVER and generalized confidence interval procedures, A Technical Report, 2011.

[19] J. Harvey, and A. J. V. Merwe, Bayesian confidence intervals for the ratio of means of lognormal data with zeros, A Microsoft word-Article, 2010.

[20] J. Harvey, and A. J. V. Merwe, Bayesian confidence intervals for means and variances of lognormal and bivariate lognormal distributions, *Journal of Statistical Planning and Inference*, *142*, 2012, 1294-1309.

[21] J. G. D'Cunha, and K. A. Rao, Bayesian inference for volatility of stock prices, *Journal of Modern Applied Statistical Methods*, *13(2)*, 2014, 493-505.

[22] J. G. D'Cunha, and K. A. Rao, Bayesian inference for mean of the lognormal distribution, *International Journal of Scientific and Research Publications*, *4(10)*, 2014,1-9.

[23] E. L. Lehmann, and J. P. Romano, *Testing Statistical Hypotheses* (Third edition) (Springer texts in Statistics, New York, 2008).