Galina NICORICI, PhD Student
E-mail: galina.nicorici@gmail.com
Faculty of Mathematics and Computer Science
University of Bucharest

OUTLIER ROBUST POSTERIOR PREDICTIVE CHECKS FOR MISSING DATA MODELS

Abstract. *The goal of this paper is to adjust the posterior predictive checks in order to overcome the problem of simultaneous outliers and missing data. We combine the approaches used to treat these two problems (separately) in the construction of particular posterior predictive p-values for Bayesian model validation. The outlier accommodation is realized using modified likelihood for estimation of posterior predictive distribution of parameter and, also, for deriving a robust form of discrepancy measure. In order to manage the missing data problem it's used the Multiple Imputation technique. According to latter, a computation algorithm for outlier robust posterior predictive p-values for missing data models is elaborated and realized in a computer environment in order to perform a simulation study, which emphasizes the efficiency and usefulness of the approach in comparison to the classical posterior predictive p-values in this particular context.*

Keywords: *Discrepancy measure, modified likelihood, multiple imputation, PCOut algorithm, posterior predictive p-value, robust diagnostics, weighted likelihood, weighting scheme.*

JEL Classification: C13, C18, C63, C88

1. Introduction

There are three ways offered by the Bayesian statistics perspective upon model checking: (1) examining sensitivity of inferences to reasonable changes in the prior distribution and the likelihood; (2) checking that the posterior inferences are reasonable, given the substantive context of the model; and (3) checking that the model fits the data. Posterior predictive checks method addresses the third of these concerns. The posterior predictive distribution is used for a *discrepancy measure*, an extension of classical test statistics designed to allow dependence on unknown

(nuisance) parameters. Sometimes it is possible to obtain a meaningful reference distribution theoretically, for the purpose of obtaining the desired tail-area probability. This is the situation of realized discrepancy assessment of model fitness proposed by Gelman et al. (1996). Otherwise, it is easily accomplished via Monte Carlo simulation. The original idea is to compare the dataset model fit to the future replications fit, in the setting that these replications are generated according to the model defined by the null hypothesis in study.

Although Bayesian methods are in general robust to some insignificant departures from initial hypothesis, potential outliers and missing data may affect the estimation and testing. As the posterior predictive checks involve both the latter, it's worth trying to protect it against breaking from outliers using a robust accommodation approach (Maronna et al. (2006)). In a Bayesian context, a natural choice is to use the modified likelihood. We try to apply this approach in the construction of our posterior predictive *p*-values. Also, to protect them from missing data, we use Multiple Imputation approach. It is convenient due to restoring of data dimensions, that's very important in the context of outliers.

In the second section of this paper we recall the definition of the posterior predictive *p*-value and describe some of its properties. The third section is about the use of modified likelihood, in Bayesian context, in order to achieve robustness in face of outliers. The fourth section is a recall of the multiple imputations technique. And the fifth section attacks straightforwardly the subject of construction of outlier robust posterior predictive *p*-values for missing data models. Afterwards a simulation example comes, comparing the robust and non-robust *p*-values in the context of contaminated and non-contaminated normal missing data models. In the end are the conclusions.

2. Posterior Predictive *P*-Values

It is obvious that, if the model fits, then the replicated data generated under the model should look similar to the observed data. To put it another way, the observed data should look plausible under the posterior predictive distribution. This is really a self-consistency check: an observed discrepancy can be due to model misfit or chance.

Our basic technique for checking the fit of a model to data is to draw simulated values from the posterior predictive distribution of replicated data and compare these samples to the observed data. Any systematic differences between the simulations and the data indicate potential failings of the model.

Let us have a dataset *y* and suppose the future replications of this one are the realizations *x* of the random variable *X*. As usual in Bayesian statistics, the parameter vector θ is also considered to be a multivariate random variable along with the data $X=(X_1,\ldots,X_n)$. Their joint posterior distribution density is defined as the following factorization:

$$f(x,\theta|y) = f(x|y,\theta)\pi(\theta|y) = f(x|\theta)\pi(\theta|y) \tag{1}$$

In this case the posterior density of the parameter, $\pi(\theta|y)$, has a natural Bayesian construction, based on the specified prior density $\pi(\theta)$ and the likelihood, represented by the *Y*'s density itself

$$\pi(\theta|y) = \frac{f(y|\theta)\pi(\theta)}{\int f(y|\theta)\pi(\theta)d\theta}. \tag{2}$$

All the above definitions are regarded in the context of testing the hypothesis $H_0 : Y \sim f(y|\theta)$ and may be conditioned on it, if it consists of more assumptions than solely the form of *Y*'s distribution. It looks like we wish to compare the unknown distribution of *Y* to the distribution proposed by the model (the known distribution of *X*) by means of some sort of loss function. For the test purpose is used a *p*-value.

Definition 1. ***Posterior predictive p-value*** *for the null hypothesis* $H_0 : Y \sim f(y|\theta)$, *against* $H_1 : Y \sim g(y|\theta), g \neq f$ *of the dataset y is the tail-area probability:*

$$p(y|H_0) = P\{D(x,\theta) \geq D(y,\theta)|y, H_0\}, \tag{3}$$

where $D(x,\theta)$ *is the discrepancy variable.* (from Meng (1994))

A *discrepancy variable* (*measure*), $D(x,\theta)$, is a scalar summary of parameters and data that is used as a standard when comparing data to predictive simulations (from de la Horra and Rodriguez-Bernal (1999)). Being a generalization of test statistics, the discrepancy variable has the same goal, to emphasize the misfit of the model to the data. Many discrepancy variables are derived from test statistics by introducing directly the dependence on the parameter, according to some properties of

the distribution density. Under some regularity conditions, a posterior predictive *p*-value which is based on a discrepancy variable arisen from a classical test statistic is asymptotically equivalent to the classical *p*-value with inserted (efficient) estimate of the parameter θ.

The probability in the definition is taken over the joint posterior distribution of (x,θ) given H_0 , defined earlier. Thus, the posterior predictive *p*-value can be regarded as the posterior mean of a classical *p*-value, averaged over the posterior distribution of the parameters under the null hypothesis. Also, Gelman et al. (1996) mention a restriction, which imposes that we can use for comparison only similar replicate samples, and that's a natural condition.

A problem claimed to affect the posterior predictive checks is that it uses the same data twice, once to fit the model parameters, and then again as input to significance test using those parameters. One of the advantages of the method is that the posterior distribution on the model parameters may well be proper, even if the prior distribution of those parameters is not. Also, using Monte Carlo techniques to work with Bayesian models tends to produce samples from the posterior distribution, making it relatively easy to estimate the posterior predictive *p*-values simply by drawing many simulated observations generated basing on each sample of parameters according to their posterior distribution, then comparing them to the real observation. The optimistic character of this measure may be seen as a disadvantage of the method. To overcome this inconvenience, Bayarri and Berger (1999) developed some sort of modifications, but the latter are not always easy to apply.

Posterior predictive checks method justifies its conclusions upon the model fit grounding on the posterior predictive *p*-values. It's almost obvious that extreme values (very close to 0 or to 1) are signs of poor fit. But, in contrast to hypothesis test procedures, here doesn't exist a definite threshold, which can assure clear rejection of a model. Rather the posterior predictive *p*-value is only a measure of fit that can lead to a more suitable model. However, for more intuitive and easier interpretation of these *p*-values Bayarri and Berger (1999) proposed a calibration of small *p*-values, which brings them closer to Bayes factor. On the other hand, the benefit of *p*-values is that they are actual posterior probabilities and, therefore, can be interpreted directly, not as posterior probability of the model being true. Moreover, Gelman and Meng (1996) state: "The role of predictive model checking is to *assess* the practical fit of a model, not to estimate the "probability that the model is true", whatever that means."

3. Modified Likelihood

The modified likelihood approach was developed in order to adjust the maximum likelihood method to obtain robustness to outliers in estimation and testing (the general method can be found in Tiku and Akkaya (2004)). This approach was expanded to be useful in Bayesian context through the use of modified likelihood function at the predictive distribution construction stage, like in Chen et al. (2000).

A common approach to outlier detection is to assess a model to explain the behaviour of the (possible) outliers. This contamination model is usually taken to be a generalization of the original model, involving an extra parameter (vector). The original model is then typically a particular case of the contamination model, corresponding to some specific value of the extra parameter(s). Testing for outliers can be reduced to testing for this specific value of the extra parameter. Conversely, someone halts at the generalization model, if the aim is the accommodation. There are plenty examples when, the particularly sensible to outliers, normal distribution is replaced by the *t* distribution to achieve robustness. But unfortunately it's not always the case that a model can be extended in an explicit way.

The likelihood function relies on the distribution density, that's why all the procedures involving it are as well affected by the outliers through the posterior density estimation. The modified likelihood approach is based on the same intuitive motivation as the idea of replacing a distribution with another one having thicker tails or a mixture of the original one with such a distribution. The idea is to directly introduce a penalization term, in order to obtain the modification of the likelihood, which would resemble to be arisen from a slightly different distribution. As well, the new form is a generalization of the classical likelihood function; hence a particular choice of extra parameter draws us back to it. Also it's natural that the departure from the original model must not be great, for we don't want to deteriorate the efficiency of inference in a non-contaminated context.

A common example of modification is the case described by Wang and Zidek (2005), namely the weighted likelihood. Its corresponding loglikelihood has a form of linear combination of logarithms of realized point probabilities:

$$\log WL(\theta, y) = \sum_{i=1}^{n} \lambda_i \log f(y_i, \theta) \tag{4}$$

Generally, fixed weights do not work very well. More reliable choice may be some adaptive weights. Our proposal in this sense is to make an adaptation of weights from the robust weighting scheme, introduced by Filzmoser et al. (2008) for outlier detection algorithm in high dimension, and to use them to obtain a robust weighted likelihood needed for posterior predictive distribution estimation.

4. Multiple Imputation

The Multiple Imputation technique was proposed by Rubin (1987) as a flexible alternative to likelihood methods for a wide variety of missing-data problems. In Multiple Imputation, each missing value is replaced by a list of $m > 1$ simulated values. Substituting the j-th element of each list for the corresponding missing value, for $j = 1, \ldots, m$, produces m plausible alternative versions of the complete data set. Each of the m data sets is analyzed in the same fashion by a complete-data method. The results are then combined by simple arithmetic, according to rules of Rubin, to obtain overall estimates and standard errors that reflect missing-data uncertainty as well as finite sample variation. This method allows the analyst to proceed with familiar complete-data techniques and software. Also, unlike other Monte Carle methods, with Multiple Imputation there is no need for a large number of repetitions for precise estimates. Generally good estimates are obtained from 5 sets, and up to 10 sets assure very good results.

5. Outlier Robust Posterior Predictive P-Values for Missing Data Models

Recall that the posterior predictive p-value is a posterior mean of a generalized p-value, averaged over the posterior distribution of parameters and (new) data under the null hypothesis. Suppressing the conditioning on H_0, we get:

$$p(y) = \int\int I_{D(x,\theta)\geq D(y,\theta)}(x) f(x|\theta)\pi(\theta|y)dxd\theta \qquad (5)$$

where I is the indicator function.

In this definition there are two quantities depending on observed y: the posterior predictive distribution of parameter, $\pi(\theta|y)$, and the discrepancy measure applied to the original data, $D(y,\theta)$. Though, these are two ways the potential

outliers may affect the posterior predictive *p*-value. Both of them need to be robustified to achieve proper accommodation.

In one of the previous sections we described the modified likelihood approach, which may be useful to achieve outlier robustness of the posterior predictive distribution of the parameter. A reliable choice is the weighted likelihood; it yields the following posterior distribution:

$$\pi(\theta|y) = \frac{\prod_{i=1}^{n} f(y_i|\theta)^{\lambda_i}\,\pi(\theta)}{\int \prod_{i=1}^{n} f(y_i|\theta)^{\lambda_i}\,\pi(\theta)d\theta}\,, \qquad (6)$$

where $\lambda_i \geq 0, i = 1,..., n$ are the chosen weights.

We stated earlier that a nice candidate for weights can be derived from Filzmoser's et al. (2008) PCOut Algorithm, which yields a robust weighting scheme of outlyingness for observations from a multivariate sample. These PCOut weights take values between 0 and 1, where outlyingness shows through values very close to 0. As the weights in the modified likelihood are introduced as powers and, in the classical likelihood, their sum equals the sample volume *n*, we want to maintain this property and rescale Filzmoser's weights accordingly,

$$\lambda_i = \frac{n w_i}{\sum_{i=1}^{n} w_i}, \quad i = 1,..., n\,, \qquad (7)$$

where w_i, $i = 1,..., n$ are the weights obtained from PCOut algorithm.

At the second step we choose a robust discrepancy variable. The natural approach is to derive a robust discrepancy variable from a simple one. Recall that a discrepancy measure is a generalization of a test statistic. In our case, for model check, a test statistic or a discrepancy variable is a function of several sufficient statistics of the sample. The idea behind the robustification is to replace classical sufficient statistics by their robust alternatives or some robust approximations, but to keep intact the relation between them and the parameters.

For our simulation example we decided to take as the original discrepancy measure the Akaike criterion, but, as the parameter dimension is fixed, it reduces to a linear function of loglikelihood. To obtain robustness, the latter is to be replaced by the same weighted likelihood:

$$D(y,\theta) = -2\log WL(\theta, y) = \sum_{i=1}^{n} \lambda_i \log f(y_i, \theta). \tag{8}$$

Having at hand both the joint posterior predictive distribution and the discrepancy variable, it's important to understand how to perform posterior predictive check in practice. As we already know, if the model is reasonably accurate, the hypothetical replications should look similar to the observed data *y*. Formally, having some realizations (obtained by means of Markov chain Monte Carlo), the necessary *p*-value is estimated as the proportion of cases in which the simulated discrepancy variable exceeds the corresponding realized value:

$$\tilde{p}(y) = \frac{1}{n}\sum_{i=1}^{n} I_{D(x^{(i)},\theta)\geq D(y,\theta)}, \tag{9}$$

where *I* is the indicator function.

Practical computation is typically a by-product of the usual Bayesian simulation that provides a set of draws of θ from the posterior distribution, $\pi(\theta|y, H_0)$. Specifically, consider the computation required for comparing the realized discrepancy $D(y,\theta)$ to its reference distribution under the null hypothesis. Given a set of (possibly dependent) draws, θ_j, j = 1, . . . , J, Gelman et al. (1996) suggest to perform the following two steps for each j:

(i) Given θ_j, draw a simulated replicated data set, x^j, from the sampling distribution, $f(x|\theta^j, H_0)$;

(ii) Calculate $D(x,\theta^j)$ and $D(y,\theta^j)$, and compare them.

Hu (1999) adjusted the former algorithm of posterior predictive checks to face properly the missing data problem by incorporating a step to reproduce the missing data pattern as a part of the simulation model and by using missing data modifications of Bayesian measures of fit. In the case of contamination hypothesis, an extra

contamination-generating step will look similar to Hu's censoring step. But, in our accommodation context settings, the contamination step is of no need, for we directly use robust measures within all computations, on both original and replicated data. We already applied with success this approach in construction of outlier robust posterior predictive *p*-values, for samples affected solely by outliers, in Nicorici (2011). In our new algorithm we also retain the censoring step proposed by Hu (1999), because the studied models imply missing data. But, for the sake of weighted robust computations, we manage the missing data estimations according to the Multiple Imputation method described previously. For every simulation step we compute new parameters basing on a set of completed selections. More in depth the procedure and the corresponding algorithm will be described in the example, for we need some model based computations developed directly from the form of model characteristic distribution, in our case the normal one.

On the other hand, we have to state explicitly the hypothesis we want to examine. For the following example we supposed that we know that the model is defined by a multivariate normal distribution. The hypothesis is then determined by the normal parameter estimates based on the observed contaminated selection with missing data. They are obtained in preliminary, according to the following algorithm:

1. Compute the estimators $\left(\hat{\mu}_0,\hat{\Sigma}_0\right)$ using the EM algorithm version adjusted to the missing data models.

2. Impute missing data according to $N\left(\hat{\mu}_0,\hat{\Sigma}_0\right)$.

3. Compute robust estimators $\left(\hat{\mu}_1,\hat{\Sigma}_1\right)$, reweighted according to the PCOut Algorithm using the previously imputed dataset.

4. Impute missing data according to the new estimators $\left(\hat{\mu}_1,\hat{\Sigma}_1\right)$.

5. Repeat steps 3-4 *m* times.

6. Combine the *m* obtained earlier estimates according to Rubin's Multiple Imputation rules.

This algorithm combines the outlier robustness adjustment approach with the Multiple Imputation technique, such that both the problems affect less each other's stages of computation. It obtains *m* sets of data needed for the Multiple Imputation

procedure, implying combining the estimates in the final step to obtain model parameter estimates to test.

6. Simulation Example

The object of this simple example is a contaminated bivariate normal sample. Its characteristic parameters are the mean vector $\mu = (\mu_1, \mu_2)' = (10,100)'$ and the covariance matrix $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} 25 & 150 \\ 150 & 2500 \end{pmatrix}$; the sample volume is $n = 50$. The data are synthetic, but their advantage is that we possess information about the underlying distribution. We will study a dataset under contamination of 10% of outliers, yield by a normal distribution having the same mean, but an "exploded" covariance matrix $\Sigma^* = 100\Sigma$. The sample is also affected by missing data, defined by three different missing data patterns, affecting only the second position of vector observations, namely: MCAR (33% missing completely at random), MAR (missing if the first component is greater than 1.12* μ_1) and MNAR (missing if the second component is greater than 1.12* μ_2).

We calculate the classical posterior predictive *p*-value to compare it to our robust posterior predictive *p*-values for both contaminated and non-contaminated samples. There are two parts of *p*-value estimation: predictive distribution estimation and discrepancy variable realizations comparison. For them both we've chosen to apply the weighted likelihood approach with PCOut weights.

Initially, the parameter $\theta = (\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$ gets the uninformative, almost flat prior with independent marginals $N(0, \sigma_0)$, $\sigma_0 = 1000$. The posterior predictive distribution of the parameter given the data then is the following:

$$\pi(\theta|y) = \frac{(2\pi)^{-\frac{5}{2}}\sigma_0^{-5}(2\pi)^{-n}}{\left(\sigma_1^2\sigma_2^2 - \rho^2\sigma_1^2\sigma_2^2\right)^{n/2}} \exp\left\{ -\frac{\mu_1^{\,2} + \mu_2^{\,2} + \sigma_1^{\,2} + \sigma_2^{\,2} + \rho^2}{2\sigma_0^2} - \right.$$

$$\left. -\sum_{i=1}^{n} \frac{(y_{i1} - \mu_1)^2\sigma_2^2 - (y_{i1} - \mu_1)(y_{i2} - \mu_2)\rho\sigma_1\sigma_2 + (y_{i2} - \mu_2)^2\sigma_1^2}{2\left(\sigma_1^2\sigma_2^2 - \rho^2\sigma_1^2\sigma_2^2\right)} \right\} \qquad (10)$$

From the latter we determine the conditionals of each component on all the others; these are needed for Markov chain Monte Carlo simulation from the joint posterior predictive distribution by means of Data Augmentation. At each step the conditional densities are reestimated and new parameters are generated by an adapted numerical form of the inverse method. Formally, the algorithm is the following:

1. Using Data Augmentation there is generated each *g*-th realization of parameter vector $(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$ according to the posterior predictive distribution from (10)

2. Given $\left(\mu_1^{(g)}, \mu_2^{(g)}, \sigma_1^{(g)}, \sigma_2^{(g)}, \rho^{(g)}\right)$ obtained in the previous step, generate $y^{(g)} = \left(y_1^{(g)}, y_2^{(g)}\right)$ from $N\left(\begin{pmatrix} \mu_1^{(g)} \\ \mu_2^{(g)} \end{pmatrix}, \begin{pmatrix} \sigma_1^{2(g)} & \rho^{(g)}\sigma_1^{(g)}\sigma_2^{(g)} \\ \rho^{(g)}\sigma_1^{(g)}\sigma_2^{(g)} & \sigma_2^{2(g)} \end{pmatrix}\right)$.

3. Apply the missing data pattern by censoring similar to the original.

4. Calculate the discordancy variables for the observed and simulated samples, $D\left(y, \theta^{(g)}\right)$ and $D\left(y^{(g)}, \theta^{(g)}\right)$, respectively.

5. Repeat the steps 1-4 many times *G* (say *G*=1000)

6. Obtain the estimated *p*-value as the mean:

$$\frac{1}{G}\sum_{g=1}^{G} I_{\left\{D\left(y^{(g)}, \theta^{(g)}\right) \geq D\left(y, \theta^{(g)}\right)\right\}},$$

where *I* is the indicator function.

The program was realized using R environment (R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.). There were used the following packages:

- *mi* – Yu-Sung Su, Andrew Gelman, Jennifer Hill, Masanao Yajima (2011). Multiple Imputation with Diagnostics (mi) in R: Opening Windows

into the Black Box. Journal of Statistical Software, 45(2), 1-31. URL http://www.jstatsoft.org/v45/i02/.

- *mnormt* – Fortran code by Alan Genz and R code by Adelchi Azzalini (2012). mnormt: The multivariate normal and t distributions. R package version 1.4-5. http://CRAN.R-project.org/package=mnormt

- *mvoutlier* – Peter Filzmoser and Moritz Gschwandtner (2012). mvoutlier: Multivariate outlier detection based on robust methods. R package version 1.9.8. http://CRAN.R-project.org/package=mvoutlier

- *norm* – Ported to R by Alvaro A. Novo. Original by Joseph L. Schafer <jls@stat.psu.edu>. (2012). norm: Analysis of multivariate normal datasets with missing values. R package version 1.0-9.4. http://CRAN.R-project.org/package=norm

The results are collected in the table below, containing the obtained *p*-values for the different settings listed above. The data are classified according to the different combinations of settings (contaminated versus non-contaminated samples with missing data, classical versus outlier and missing data robust *p*-values).

Table 1

Classic versus outlier robust posterior predictive *p*-values for a sample from a particular non-contaminated and contaminated bivariate normal with missing data

Missing pattern / **P-value type**	***Non-contaminated sample***		
	MCAR	MAR	MNAR
Classic	0.2935	0.0465	0.3505
Weighted robust	0.3950	0.2837	0.4137
	Contaminated sample		
Classic	0.2310	0.0000	0.0000
Weighted robust	0.3068	0.2642	0.3810

We see that robust *p*-values perform pretty well in confirming the real hypothesis on non-contaminated samples, it means that they do not lose much in efficiency comparing to the classical *p*-values. The classical posterior predictive *p*-value breaks in the sensible normal contaminated sample with missing data.

Repeated simulations with plenty different input parameters and context settings lead to similar results. So we conclude that the approach was successful.

6. Conclusions

In this paper it was proposed an approach to build robust posterior predictive *p*-values to protect the posterior predictive checks method from the influence of both potential outliers and missing data. The robust procedure was realized in software environment and tested on a simulated normal sample. Future research may imply the study of other distribution hypotheses. Also it would be useful a research implying construction of other robust discrepancy variable to test some specific properties of the null hypothesis model, for more sensible analysis.

REFERENCES

[1] **Bayarri, M.J., Berger, J.O. (1999),** ***Quantifying Surprise in the Data and Model Verification***. In *Bayesian Statistics* 6 (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.) p. 53-82. *Oxford Univ. Press*;

[2] **Chen, Y., Breen, P.A., Andrew, N.L. (2000),** ***Impacts of Outliers and Misspecification of Priors on Bayesian Fisheries-stock Assessment***. *Can. J Aquat. Sci*., v. 57, p. 2293–2305;

[3]**Filzmoser, P. , Maronna, R., Werner, M. (2008),** ***Outlier Identification in High Dimensions*** .*Comp. Stat. & Data Analysis*, v. 52, n. 3, p. 1694-1711;

[4]**Gelman, A., Meng, X. (1996),** ***Model Checking and Model Improvement***. (A chapter for Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. (Eds.) *Markov chain Monte Carlo in Practice*. Chapman and Hall, London, UK;

[5]**Gelman, A., Meng, X., Stern, H. (1996),** ***Posterior Predictive Assessment of Model Fitness via Realized Discrepancies.*** *Statistica Sinica*, v. 6, p. 733-807;

[6]**Horra, J., Rodriguez-Bernal, M. (1999),** ***The Posterior Predictive P-Value: An Alternative to the Classical P-Value*** . *Rev. R. Acad. Cienc. Exact. Fis. Nat*, v. 93, n. 3, p. 321-328;

[7]**Hu, M. (1999),** ***Model Checking for Incomplete High Dimensional Categorical Data***. Ph.D. Thesis, University of California, Los Angeles;

[8]**Maronna, R. A., Martin, R. D., Yohai, V. J. (2006),** ***Robust Statistics Theory and Methods***. Wiley;
[9]**Meng, X. (1994),** ***Posterior Predictive P-Values***. *Ann. Statist.*, v. 22, n. 3, p. 1142-1160;
[10]**Nicorici, G. (2011),** ***Outlier Robust Posterior Predictive Checks***. *Analele Universitatii Bucuresti*, seria Matematica-Informatica, Anul LX-2011, p. 69—82;
[11]**Rubin, D. B. (1987), Multiple Imputation for Non-response in Surveys**. *New York: Wiley;*
[12]**Tiku, M. L., Akkaya, A. D. (2004),** ***Robust Estimation and Hypothesis Testing*****.** New Delhi: New Age International (formerly Wiley Eastern);
[13]**Wang, X., Zidek, J. (2005),** ***Selecting Likelihood Weights by Cross-Validation*****.** *Ann. Statist.*, v. 33, n. 2, p. 463-500.