Center for
Economic Theories and Policies
Sofia University St. Kliment Ohridski
Faculty of Economics and Business Administration

ISSN: 2367-7082

Bulgarian Economic Papers

Bayesian Forecast Intervals for Inflation and Unemployment Rate in Romania

Mihaela Simionescu

BEP 06-2017
Publication: May 2017

Bayesian forecast intervals for inflation and unemployment rate in Romania

Mihaela Simionescu[1]

May 28, 2017

Abstract: This paper brings as novelty for the Romanian literature the construction of Bayesian forecast intervals for inflation and unemployment rate in the period 2004-2017. Only few intervals included the registered values on the variables, but in the last stage when all the prior information has been used, the forecast intervals are very short. On the other hand, a novelty for the international literature is brought in this research by proposing a Bayesian technique for assessing prediction intervals in a better way than in the traditional approach that uses statistic tests.

Key-words: forecast interval, Bayesian interval, inflation, unemployment

JEL: C11, C13, C53, E37

[1] Institute for Economic Forecasting, mihaela_mb1@yahoo.com

1. Introduction

The main aim of this paper is to build forecast intervals for inflation and unemployment rate by using the Bayesian method. Moreover, an alternative method to Christoffersen (1998, p. 842) tests to assess the forecast intervals is proposed by using a Bayesian approach. In Romania, few papers were dedicated to prediction intervals, even if this is a priority for National Bank of Romania in the context of inflation targeting and uncertainty evaluation.

The Bayesian forecast intervals consider different degrees of uncertainty. A lower degree of uncertainty is associated with a shorter interval, but there are high chances for the registered value to be out of the prediction interval.

Forecast intervals help policy-makers in selecting the future macroeconomic policy, but taking into consideration that inflation or unemployment is located in a certain interval. For National Bank of Romania some uncertainty intervals are built for the inflation rate in the context of this indicator targeting, but the intervals are based on the extrapolation of the mean errors corresponding to past projections. On the other hand, it would be useful for National Bank of Romania to use a fan chart in reporting the degree of uncertainty corresponding to their predictions. This fan chart is based on the calculation of some objective probabilities.

The paper continues as follows. After the methodological background of the literature realizations, Bayesian forecast intervals are built and the results from a sample are processed in order to evaluate the intervals. The last section brings conclusions.

2. Assessing forecast intervals

Information about predictions' accuracy is not reflected by point predictions (Clements, 2014, p. 208). Therefore, it is useful to use forecast intervals that permit the quantification of future uncertainty. Moreover, other advantages of forecast intervals are: prediction methods comparisons and establishment of desired strategies to get a certain target.

For Romania, Bratu (2012, p. 146) and Simionescu (2015, p. 90) built some forecast intervals for unemployment and inflation rate. The historical errors method was applied by keeping constant a certain error measure. On the other hand, another method for constructing prediction intervals is bootstrap method. It is used when the sample' distribution is not known.

An alternative to bootstrap method is the grid bootstrap employed by Gospodinov (2002, p. 87) for calculating the median without any bias. In this case, there should be a high persistence of the modeled variable. However, this methods has as disadvantage the intensive computation (Guan, 2003, p. 79). Consistent estimators for the conditional distribution are considered in the intervals based on sieve bootstrap. It is a non-parametric method used by Alonso, Pena and Romo (2003, p. 185).

Surprisingly, in literature there are only few studies that were interested in building Bayesian prediction intervals. Hamada et al. (2004, p. 454) proposed Bayesian forecast intervals that include a part from a finite number of values with a specified probability. A comparison of these intervals with tolerance intervals is made. Smith et al. (2009, p. 215) proposed an algorithm for estimating forecast intervals based on Bayesian predictions model. Wallis (2003, p. 26) assessed forecast intervals by making inferences on p-values. Christoffersen (1998, p. 34) described various coverage tests used in assessing the forecast intervals (conditional, unconditional and independent coverage tests).

I. Likelihood ratio (LR) tests

Let us consider a chronological series for the prediction intervals. Let us establish a probability of π that the value be inside that interval. Time series for registered and predicted values are considered and the aim is to check if the ex-ante probability that registered value be in the prediction interval is correct. A number of n_1 values are in the forecast intervals and a number of n_2 intervals are outside. The coverage probability represents a ratio: p=n_1/n. we have a binomial distribution. Under the null hypothesis, the likelihood is computed as: $L(\pi) = (1-\pi)^{n_2} \pi^{n_1}$. Under the alternative hypothesis likelihood becomes: $L(p)=(1-p)^{n_2}\,p^{n_1}$. Likelihood ratio statistic is: $LR_{UC}=2(n_2\log\frac{1-p}{1-\pi}+n_1\log\frac{p}{\pi})\underset{H_0}{\longrightarrow}\chi_1^2$.

This is the unconditional coverage test, which is rather unsuitable for chronologic series. Therefore, another test was proposed by Christoffersen (1998, p. 845) by combining the unconditional coverage test with the independence test.

The independence test uses the matrix of transition frequencies $[n_{ij}]$ that represents the number of values in state i at time (t-1) and in stat j at time t. Maximal likelihood estimations are considered for transition probabilities. It is a ratio between frequencies in a cell and total frequencies of a line. Two situations are taken into consideration: the value is inside or it is outside the interval, which corresponds to 1 and 0. The transition matrix for these estimated probabilities is:

$$P=\begin{pmatrix}1-p_{01} & p_{01}\\ 1-p_{11} & p_{11}\end{pmatrix}=\begin{pmatrix}\frac{n_{00}}{n_{0.}} & \frac{n_{01}}{n_{0}\cdot}\\ \frac{n_{10}}{n_{1.}} & \frac{n_{11}}{n_{1.}}\end{pmatrix}$$

The likelihood associated to P is: $L(P)=(1-p_{01})^{n_{00}}\cdot p_{01}^{n_{01}}\cdot(1-p_{11})^{n_{10}}\cdot p_{11}^{n_{11}}$.

The null hypothesis for the independence test states that the (t-1) and t states are independent, which means that $\pi_{01}=\pi_{11}$.

The maximal likelihood estimator for joint probability is: $p=\frac{n_{.1}}{n}$. The likelihood under the null hypothesis evaluated for the value of p is: $L(p)=(1-p)^{n_{.0}}\cdot p^{n_{1.}}$.

The LR test statistic will be: $LR_{ind}=-2\log\frac{L(p)}{L(P)}\underset{H_0}{\longrightarrow}\chi_1^2$. Christoffersen (1998, p. 850) combined the unconditional coverage test with independence test resulting: $LR_{CC}=-2\log\frac{L(\pi)}{L(P)}\underset{H_0}{\longrightarrow}\chi_2^2$.

In the case that the first observation is eliminated, we have : $LR_{CC}=LR_{UC}+LR_{ind}$.

II. Chi-square (χ^2) tests

Likelihood ratio tests are equivalent, in statistic terms, with Pearson's goodness-of-fit tests. Wallis (2008, p. 20) utilized them for assessing forecast intervals.

The Chi-square test for unconditional coverage is based on: $X^2 = \frac{n(p-\pi)^2}{\pi(1-\pi)}$.

Let the matrix of observed frequencies be $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. In this case, the test statistic is:

$X^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$.

The conditional coverage test linked with the independence test is based on contingency table for observed frequencies and expected frequencies. The null hypothesis states independent lines and the coverage probability is π. The expected frequencies matrix is: $\begin{bmatrix} (1-\pi)(a+b) & \pi(a+b) \\ (1-\pi)(c+d) & \pi(c+d) \end{bmatrix}$. The test has 2 degrees of freedom. The statistic is a sum of square normal standard statistics for samples proportions. For each table's row, a proportion is given. For small samples, the sum of LR statistics cannot be transposed in chi-square test terms. In the proposed method for forecast interval assessment I will use the Bayesian approach and I will compute a probability for each prediction interval.

3. Assessing forecast intervals for inflation and unemployment rate in Romania

The forecast intervals might be constructed using the Bayesian approach that allows for different degrees of uncertainty that are caused by the quantity of available data. We start from the data for inflation and unemployment rate that were provided by 4 experts in forecasting. The forecasts' averages, the selection variances and the intervals for each year are computed. A level of significance of 5% is considered. According to Bayesian approach, prior information should increase the estimations' accuracy. In the case of intervals' construction, the Bayesian approach supposes getting shorter intervals. In the first stage, the case of total uncertainty is considered. Then, we take into consideration the predictions of first two forecasters for which mean, variance and intervals are calculated. In the next stages, the predictions of the third and the fourth provider are taken into consideration.

The revised probability density is calculated using the information about average. A weighted average is computed using as weights the inverse of variances. That information with lower dispersion is more accurate and it will receive a larger weight. The last interval includes all the information regarding predictions for the considered variables.

Table 1. ***Bayesian forecast intervals for inflation rate in Romania (horizon: 2004-2017)***

Year	Selection mean	Selection dispersion	Bayesian forecast interval	
Stage 1				
2004	12.75	4.315	8.67	16.82
2005	9.29	0.384692	8.08	10.51
2006	7.08	0.411225	5.82	8.33
2007	5.13	2.718892	1.90	8.36
2008	4.53	2.773267	1.26	7.79
2009	3.56	2.282025	0.60	6.52
2010	3.88	4.743267	-0.38	8.14
2011	4.55	0.376667	3.34	5.75
2012	5.27	4.229167	1.24	9.30
2013	3.77	1.595833	1.29	6.25
2014	1.13	1.266233	-1.07	3.34
2015	1.02	1.225853	-0.58	3.15
2016	0.92	1.027453	-0.76	2.29
2017	1.23	1.936237	0.92	1.87

Year	Selection mean	Selection dispersion	Bayesian forecast interval	
Stage 2				
2004	14.05	5.78	9.33	18.76
2005	9.09	0.49005	7.72	10.46
2006	6.66	0.03645	6.29	7.03
2007	4.01	0.03645	3.64	4.38
2008	5.16	6.5522	0.14	10.17
2009	2.37	1.08045	0.33	4.41
2010	2.06	0.4802	0.70	3.41
2011	5.05	0.005	4.91	5.18
2012	4.55	0.005	4.41	4.68
2013	4.75	0.005	4.61	4.88
2014	0.22	0.2048	-0.66	1.10
2015	0.78	0.0186	-0.22	2.56
2016	0.55	0.2925	0.16	2.46
2017	1.01	0.3438	1.03	1.57
Stage 3				
2004	13.33	4.430833	9.20	17.45
2005	9.06	0.248033	8.08	10.03
2006	6.77	0.055633	6.31	7.23
2007	4.34	0.341633	3.19	5.48
2008	4.64	4.0873	0.67	8.60
2009	3.08	2.045433	0.28	5.88
2010	3.44	5.9533	-1.34	8.22
2011	4.63	0.523333	3.21	6.05
2012	4.26	0.243333	3.29	5.23
2013	4.33	0.523333	2.91	5.75
2014	0.71	0.832533	-1.07	2.50
2015	1.02	0.342856	-0.10	2.04
2016	0.75	0.337697	0.29	1.04
2017	1.13	0.337524	0.97	1.14
Stage 4				
2004	13.69	0.063458	13.19	14.18
2005	9.07	0.065753	8.57	9.58
2006	6.72	0.067797	6.21	7.23
2007	4.17	0.070093	3.66	4.69
2008	4.90	0.071301	**4.37**	**5.42**
2009	2.72	0.071514	2.20	3.25
2010	2.75	0.07486	2.21	3.28
2011	4.84	0.069767	4.32	5.35
2012	4.40	0.070755	3.88	4.92
2013	4.54	0.085227	3.96	5.11
2014	0.46	0.070012	-0.05	0.98
2015	0.67	0.056390	0.05	0.87
2016	0.48	0.447939	0.38	0.76
2017	0.67	0.227845	0.44	0.87

For computing the intervals in the last stage, some weights are used (w1 and w2). These weights are computed as inverse of variances from the previous two stages. The mean is computed as arithmetic average of the means in the previous two stages, while variance is the inverse of the sum of weights.

The results indicated that the intervals from the last stage are shorter than those based on prior information or on experts' opinions. However, only one interval (that for 2008) included the registered one.

Table 2. ***Bayesian forecast intervals for unemployment rate in Romania (horizon: 2004-2017)***

Year	Selection mean	Selection dispersion	Bayesian forecast interval	
Stage 1				
2004	7.97	0.054692	7.51	8.43
2005	7.74	0.096358	7.13	8.35
2006	7.49	0.130967	6.78	8.20
2007	7.34	0.2271	6.41	8.27
2008	7.19	0.318225	6.08	8.29
2009	7.43	1.032133	5.43	9.42
2010	7.11	0.923225	5.23	9.00
2011	7.25	0.043333	6.84	7.65
2012	7.1	0.046667	6.67	7.52
2013	6.32	1.0425	4.32	8.32
2014	7.03	0.039033	6.64	7.42
2015	6.9	0.036577	6.75	7.25
2016	6.56	0.046578	6.45	7.20
2017	6.04	0.559754	5.98	6.34
Stage 2				
2004	7.85	0.08405	7.28	8.01
2005	7.54	0.12005	6.86	7.78
2006	7.29	0.2048	6.40	7.69
2007	7.04	0.3042	5.95	7.63
2008	6.93	0.68445	5.31	8.27
2009	6.71	0.8978	4.85	8.46
2010	6.53	1.23245	4.35	8.95
2011	7.1	0.02	6.82	7.13
2012	7	0.02	6.72	7.03
2013	5.6	0.98	3.65	7.52
2014	7.17	0.0242	6.86	7.21
2015	7.01	0.2294	6.87	7.13
2016	6.87	0.3673	6.67	7.12
2017	6.45	0.4386	6.05	6.28

Year	Selection mean	Selection dispersion	Bayesian forecast interval	
Stage 3				
2004	7.90	0.049033	7.46	8.33
2005	7.66	0.102033	7.03	8.28
2006	7.46	0.1891	6.60	8.31
2007	7.22	0.256633	6.23	8.21
2008	7.09	0.4143	5.82	8.35
2009	7.27	1.400933	4.95	9.59
2010	6.82	0.865633	4.99	8.64
2011	7.23	0.063333	6.74	7.72
2012	7.13	0.063333	6.64	7.62
2013	6.13	1.343333	3.86	8.40
2014	7.11	0.021733	6.82	7.40
2015	6.98	1.0384656	6.97	7.11
2016	6.56	0.6573854	6.50	6.78
2017	6.34	1.0364655	6.24	6.40
Stage 4				
2004	5.10	7.514089	0.26	10.47
2005	0.36	4.502814	3.79	4.52
2006	0.04	2.538715	3.07	3.16
2007	0.18	1.783061	2.42	2.80
2008	5.31	0.910125	**3.44**	**7.18**
2009	1.56	0.435022	0.27	2.85
2010	3.21	0.476625	1.86	4.56
2011	0.26	12	**6.52**	**7.05**
2012	0.12	12	**6.66**	**6.91**
2013	0.26	0.430416	1.02	1.55
2014	0.51	21.77068	8.62	9.66
2015	3.23	10.39444	3.15	3.25
2016	2.06	5.485639	2.01	2.11
2017	1.04	4.346553	1.00	1.15

Even if the intervals' length in the last stage is lower, only several intervals included the registered unemployment rate (forecasts' intervals in 2008, 2011 and 2012). 20 forecasters were randomly selected. They provided forecast intervals for inflation rate in 2016. After the end of 2016, when the registered value of inflation was obtained, five experts were randomly selected. We checked if they provided a forecast intervals that included the registered inflation rate. Three out of the five selected experts provided prediction intervals that included the registered inflation rate in 2016. The aim is to estimate the number of experts that offered good forecast intervals in the entire population of 20 forecasters.

The estimator is seen as a confidence distribution of the possible values. Before selecting the 5 forecasters, no information was available about the total number of experts that provided correct prediction intervals. For having a certain prior information, we make the assumption that half of the experts provided intervals that included inflation rate. In this particular case, we have a Binomial distribution: Binomial (50%, 20). For computing it, we will use the Excel function, if q is the number of people that provided the correct interval in the entire population: p(q) =BINOMDIST(q,20,0.5,0).

The likelihood function is related to the sample of 5 experts from population follow hypergeometric distribution, where we know: total population volum (M = 20), sample volum (n = 5), number of experts that fulfill the required condition (x = 3). We do not know the number of experts in the total population of 20 people that offert correct intervals (q). The parameter that should be estimated is denoted by q. *Hypgeomdist* function from Excel is used to compute the likelihood function.

I(X|q) =HYPGEOMDIST(3,5,q,20).

Tabel 3. ***Estimation of q***

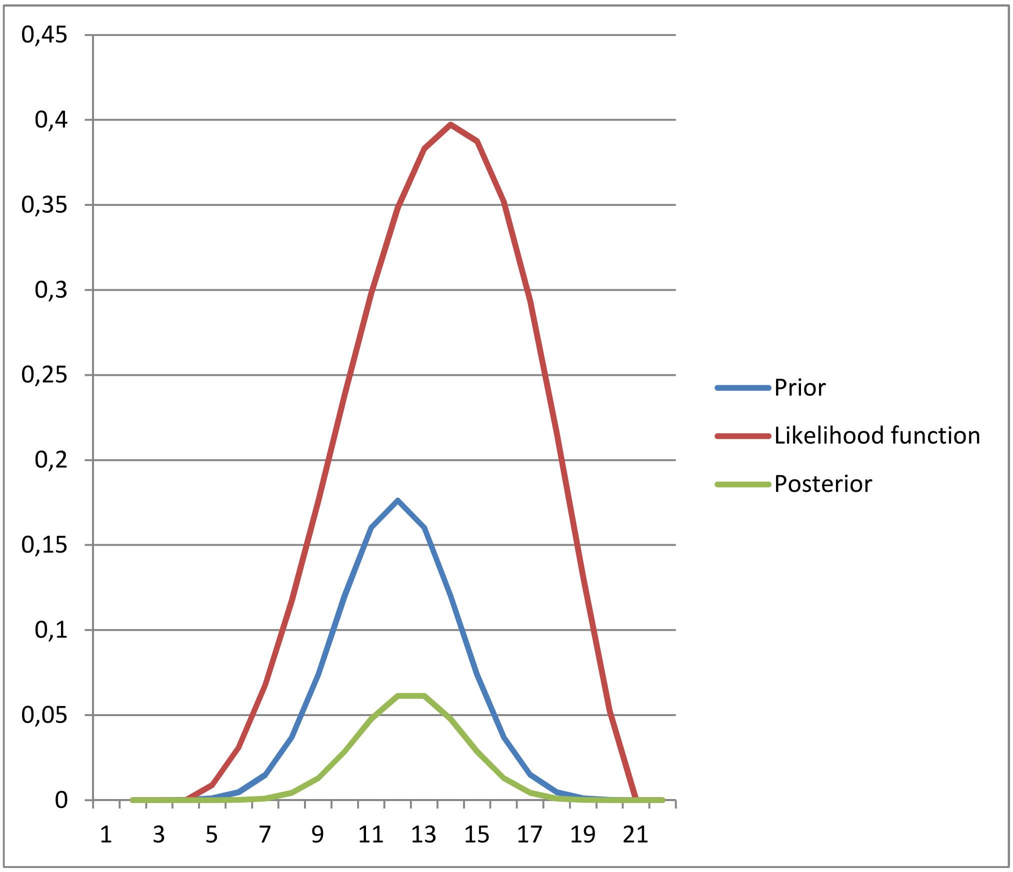
q	Prior	Likelihood function	Posterior	Normalized posterior
0	9.5E-07	0	0.0E+00	0.0E+00
1	1.9E-05	0	0	0
2	1.8E-04	0	0	0
3	1.1E-03	8.8E-03	9.5E-06	3.1E-05
4	4.6E-03	3.1E-02	1.4E-04	4.6E-04
5	1.5E-02	6.8E-02	1.0E-03	3.2E-03
6	3.7E-02	1.2E-01	4.3E-03	1.4E-02
7	7.4E-02	1.8E-01	1.3E-02	4.2E-02
8	1.2E-01	2.4E-01	2.9E-02	9.2E-02
9	1.6E-01	3.0E-01	4.8E-02	1.5E-01
10	1.8E-01	3.5E-01	6.1E-02	2.0E-01
11	1.6E-01	3.8E-01	6.1E-02	2.0E-01
12	1.2E-01	4.0E-01	4.8E-02	1.5E-01
13	7.4E-02	3.9E-01	2.9E-02	9.2E-02
14	3.7E-02	3.5E-01	1.3E-02	4.2E-02
15	1.5E-02	2.9E-01	4.3E-03	1.4E-02
16	4.6E-03	2.2E-01	1.0E-03	3.2E-03
17	1.1E-03	1.3E-01	1.4E-04	4.6E-04
18	1.8E-04	5.3E-02	9.5E-06	3.1E-05
19	1.9E-05	0	0	0
20	9.5E-07	0	0	0

For computing the likelihood function, we will consider that the estimation q cannot be 0,1, or 2, because we identified in the sample a number of 3 experts that provided correct forecast intervals. On the other hand, q cannot be 19 or 20, because 2 experts were identified that they do not provide a suitable interval.

Posterior value is computed by multiplying prior with likelihood function. The normalized posterior value is computed by dividing each posterior value to sum of posterior values.

The peak of posterior is under the peak of prior distribution and likelihood function. The impact of the likelihood function was quite low, but the prior had a higher effect of posterior. The effect of likelihood function is low, because the selected sample was the low volume. The prior has a maximum value when q is 10.

Fig. 1. ***Prior distribution, posterior distribution and likelihood for q***



4. Conclusions

In this paper, some Bayesian forecast intervals are proposed for inflation and unemployment rate of Romania. Compared to previous studies on prediction intervals, we can state that this research is the first proposal of this type of intervals for Romanian literature. On the other hand, we also proposed an alternative way of assessing forecast intervals by taking into consideration the Bayesian approach that links prior information to sample information. This method is very useful for evaluating intervals when we have only few experts that provided prediction intervals and we are interested in making generalizations for the entire population of forecasters in that country.

The results indicated that the successive incorporation of prior information diminishes the forecast intervals length, but the major disadvantage is that only few intervals included the registered value for inflation, respectively unemployment rate. We considered that the Bayesian approach for assessing some forecast intervals is better than the traditional approach, because a precise probability is computed for each interval that might include the registered value.

In a future research it is important to propose other forecast intervals based on other methods. A comparison between those forecast intervals with the Bayesian approach using the proposed method in this paper would be very useful in order to detect the best prediction intervals. These suitable intervals based on a certain method would be necessary for decision making process at macroeconomic level.

References

Alonso, Andreas, Daniela Pena and Juan Romo, J. 2000.''Sieve Bootstrap Prediction Intervals.'' *Proceedings In Computational Statistics 14th Symposium* 1: 181-186.

Bratu, Mihaela. 2012." Forecast Intervals for Inflation in Romania." *Timisoara Journal of Economics 5*(1(17)): 145-152.

Christoffersen, Peter. 1998. "Evaluating Interval Forecasts." *International Economic Review* 39(4): 841-862.

Clements, Michael. 2014. "Forecast Uncertainty—Ex Ante and Ex Post: US Inflation and Output Growth." *Journal of Business & Economic Statistics 32*(2): 206-216.

Gospodinov, Nikolay. 2002. "Median unbiased forecasts for highly persistent autoregressive processes." *Journal of Econometrics* 111(1): 85-101.

Guan, Weihua. 2003. "From the help desk: bootstrapped standard errors." *The Stata Journal* 3(1): 71–80.

Hamada, M.ichael, Johnson, Valen, Moore, Leslie M., Wendelberger, Joanne. 2004. "Bayesian prediction intervals and their relationship to tolerance intervals." *Technometrics 46*(4): 452-459.

Simionescu, Miahela. 2015. "The assessment of some point and forecast intervals for unemployment rate in Romania." *International Journal of Economic Practices and Theories 5*(2): 88-94.

Smith, Amber D., Genz, Alan, Freiberger, David M., Belenky, Gregory, Van Dongen, Hans P. 2009. "Efficient computation of confidence intervals for Bayesian model predictions based on multidimensional parameter space." *Methods in enzymology 454*: 213-231.

Wallis, Kenneth. 2008. "Forecast uncertainty, its representation and evaluation." *Econometric Forecasting and High-Frequency Data Analysis 13*: 1-51.