DOI: 10.5604/01.3001.0014.0532

PRZEGLĄD STATYSTYCZNY
TOM LXV – ZESZYT 2 – 2018

Anna STASZEWSKA-BYSTROVA[1]

Refined Bonferroni prediction bands for autoregressive models[2]

1. INTRODUCTION

A number of papers address the problem of constructing prediction bands for multivariate and univariate autoregressive (AR) models (see e.g., Jordà, Marcellino, 2010; Staszewska-Bystrova, 2011, 2013; Staszewska-Bystrova, Winker, 2013; Wolf, Wunderli, 2015). Joint prediction bands are designed to contain the future trajectory of a predicted variable with probability given by the coverage level and therefore provide valuable information on the predictive uncertainty. The most successful methods of band construction use the bootstrap (Efron, 1979) to derive the relevant predictive distributions. Bootstrap methods are also commonly used for this class of models for forming prediction intervals (see i.a. Thombs, Schucany, 1990; Masarotto, 1990; Breidt et al., 1995; Grigoletto, 1998; Kim, 2001; Clements, Kim, 2007).

The methods of building joint bands, which have been proposed, lead to obtaining prediction regions which differ with respect to the estimated coverage levels and widths. Simulation studies reported by Lütkepohl et al. (2015*a*, 2015*b*) in the context of constructing confidence bands for impulse responses show that the conservative bootstrap Bonferroni bands are quite successful in terms of maintaining the nominal coverage probability. However, the estimated coverage rates are often larger than the nominal values for these bands. Excessive coverage is, in turn, associated with unnecessarily large width of the bands.

The aim of this paper is to refine the basic bootstrap Bonferroni bands in two ways: first, by applying higher order Bonferroni-type inequalities (Hoover, 1990, see also Glaz, Ravishanker, 1991) and second, by considering imbalanced Bonferroni bands found through optimization. Both refinements should lead to reductions in the width of the bands. The working of the methods is compared to the performance of the sup-*t* procedure described by Wolf, Wunderli (2015).

[1] University of Lodz, Faculty of Economics and Sociology, Chair of Econometric Models and Forecasts, 41 Rewolucji 1905r. St., 90–214 Lodz, Poland, e-mail: anna.bystrova@uni.lodz.pl.

[2] Financial support provided by the National Science Center, Poland (NCN) through HARMONIA 6: UMO-2014/14/M/HS4/00901 is gratefully acknowledged.

The proposed methods are applied to persistent AR models containing a linear trend. The parameters of such models can be estimated using various methods. Some standard estimators including the ordinary least squares (OLS) or the Yule-Walker estimators are, however, not recommended due to their small sample bias (see e.g. Andrews, Chen, 1994). Alternative estimation methods for univariate AR models with time trend have been proposed for instance by Andrews, Chen (1994), Kilian (1998) and Roy, Fuller (2001). Clements, Kim (2007) report that bootstrap prediction intervals for the AR model based on the approximately median unbiased Roy-Fuller estimator have the best small sample properties. This estimator is therefore applied in the study reported below.

The structure of the paper is as follows. The next section presents the AR framework and the estimation method used. In section 3 the bootstrap algorithm for obtaining predictive distributions is described and in section 4 the standard Bonferroni bands, the proposed refinements and the benchmark sup-*t* bands are discussed. Section 5 presents the Monte Carlo comparison of the methods while section 6 concludes.

2. THE MODEL

The model considered in this paper is an AR(*p*) with intercept and a linear time trend (see Box, Jenkins, 1970; Lütkepohl, Krätzing, 2004):

$$y_t = \mu + \beta t + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p} + \varepsilon_t, \tag{1}$$

where $\varepsilon_t \sim \mathrm{iid}(0, \sigma^2)$.

The model can be reparametrized either as

$$y_t = \mu + \beta t + \gamma_1 y_{t-1} + \delta_1 \Delta y_{t-1} + \cdots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t, \tag{2}$$

where $\Delta y_t = y_t - y_{t-1}$, $\gamma_1 = \sum_{i=1}^{p} \alpha_i$, $\alpha_1 = \gamma_1 + \delta_1$, $\alpha_i = \delta_i - \delta_{i-1}$ for $2 \leq i \leq p-1$ and $\alpha_p = -\delta_{p-1}$ or as

$$y_t = \mu + \beta t + \gamma_{-1} y_{t-1} + \theta_1 S y_{t-1} + \cdots + \theta_{p-1} S y_{t-p+1} + \varepsilon_t, \tag{3}$$

where $S y_t = y_t + y_{t-1}$, $\gamma_{-1} = \sum_{i=1}^{p} (-1)^{i+1} \alpha_i$, $\alpha_1 = \gamma_{-1} + \theta_1$, $\alpha_i = \theta_i + \theta_{i-1}$ for $2 \leq i \leq p-1$ and $\alpha_p = \theta_{p-1}$. The parameter γ_1 in (2) describes the persistence of the AR process. In what follows it is assumed to belong to the interval $(-1,1)$.

Given the pre-sample values $y_{-p+1}, \ldots, y_0$ and the sample values $y_1, \ldots, y_T$, the parameters of model (2) can be estimated using the method proposed by Roy,

Fuller (2001). The Roy-Fuller (RF) estimator of γ_1 is approximately unbiased and its mean squared error is smaller than that of the ordinary least squares estimator for time series with a root near 1. The estimator is defined as

$$\hat{\gamma}_1^{RF} = \min(\tilde{\gamma}_1, 1), \tag{4}$$

where $\tilde{\gamma}_1 = \dot{\gamma}_1 + [C_p(\dot{\tau}_1) + C_{-p}(\dot{\tau}_{-1})]\dot{\sigma}_1$, $\dot{\gamma}_1$ is the least squares estimator of the parameter of $\hat{y}_{t-1}$ in the model where $\hat{y}_t$ is regressed on $\hat{y}_{t-1}, \Delta\hat{y}_{t-1}, \ldots, \Delta\hat{y}_{t-p+1}$, $\hat{y}_t$ denotes the least squares residual from the regression of y_t on the constant and linear trend t and $\dot{\sigma}_1$ is the standard error of $\dot{\gamma}_1$. The functions $C_p(\dot{\tau}_1)$ and $C_{-p}(\dot{\tau}_{-1})$ are based respectively on the unit root statistic

$$\dot{\tau}_1 = \frac{\dot{\gamma}_1 - 1}{\dot{\sigma}_1}, \tag{5}$$

and the negative unit-root statistic

$$\dot{\tau}_{-1} = \frac{\dot{\gamma}_{-1} + 1}{\dot{\sigma}_{-1}}, \tag{6}$$

where $\dot{\gamma}_{-1}$ and $\dot{\sigma}_{-1}$ are the least squares estimator of the coefficient of $\hat{y}_{t-1}$ in the model where $\hat{y}_t$ is regressed on $\hat{y}_{t-1}, S\hat{y}_{t-1}, \ldots, S\hat{y}_{t-p+1}$ and the standard error of $\dot{\gamma}_{-1}$. $C_p(\dot{\tau}_1)$ has the form[3]

$$\begin{aligned}
C_p(\dot{\tau}_1) &= -\tau_{med} + d_T(\dot{\tau}_1 - \tau_{med}), && \dot{\tau}_1 > \tau_{med},\\
&= I_p(T^{-1}\dot{\tau}_1) - 3[\dot{\tau}_1 + k(\dot{\tau}_1 - K)]^{-1}, && K < \dot{\tau}_1 \le \tau_{med},\\
&= I_p(T^{-1}\dot{\tau}_1) - 3[\dot{\tau}_1]^{-1}, && -\sqrt{\frac{3T}{I_p}} < \dot{\tau}_1 \le K,\\
&= 0, && \dot{\tau}_1 \le -\sqrt{\frac{3T}{I_p}},
\end{aligned}$$

where I_p stands for the integer part of $\frac{1}{2}(p+1)$, τ_{med} is the median of the limiting distribution of $\dot{\tau}_1$ under the null, $k = \left[3T - \tau_{med}^2\left(I_p + T\right)\right][\tau_{med}(\tau_{med} - K)(I_p + +T)]^{-1}$. The constants *K* and d_T are set, following Roy, Fuller (2001), to –5 and 0.29, respectively. The expression for $C_{-p}(\dot{\tau}_{-1})$ is as follows

[3] Apart from the paper by Roy, Fuller (2001) see also the errata available at Anindya Roy's webpage.

$$\begin{aligned} C_{-p}(\dot{t}_{-1}) &= 0, && \dot{t}_{-1} \geq \sqrt{k_{-1}}, \\ &= \left(\left\lfloor \frac{p+1}{2} \right\rfloor + I_{2p}\right) T^{-1}\dot{t}_{-1} - \dot{t}_{-1}^{-1}, && K \leq \dot{t}_{-1} < \sqrt{k_{-1}}, \\ &= a_T + b_T(\dot{t}_{-1} + K), && \dot{t}_{-1} < K, \end{aligned}$$

where I_{2p} is equal to 0 if p is even and is given by 3 if p is odd, $\lfloor q \rfloor$ stands for the greatest integer less than or equal to q, $k_{-1} = \left(\left\lfloor \frac{p+1}{2} \right\rfloor + I_{2p}\right)^{-1} T$, $a_T = C_{-p}^{*}(-K)$, $b_T = C_{-p}^{*\prime}(-K)$, and $C_{-p}^{*}(\dot{t}_{-1}) = \left(\lfloor (p+1)/2 \rfloor + I_{2p}\right) T^{-1}\dot{t}_{-1} - \dot{t}_{-1}^{-1}$.

Given an estimate $\hat{\gamma}_1^{RF}$, the parameters $\mu, \beta, \delta_1, \ldots, \delta_{p-1}$ can be estimated from the regression of $y_t - \hat{\gamma}_1^{RF} y_{t-1}$ on the constant, trend and lagged differences $\Delta y_{t-1}, \ldots, \Delta y_{t-p+1}$, producing $\hat{\mu}^{RF}, \hat{\beta}^{RF}, \hat{\delta}_1^{RF}, \ldots, \hat{\delta}_{p-1}^{RF}$. The only exception arises if $\hat{\gamma}_1^{RF} = 1$ when the parameter on trend is restricted to 0. In the next step, estimates of $\alpha_1, \ldots, \alpha_p$, denoted by $\hat{\alpha}_1^{RF}, \ldots, \hat{\alpha}_p^{RF}$, can be obtained. The variance of the random error ε_t can be estimated using

$$\hat{\sigma}^2 = \frac{1}{T-l} \sum_{t=1}^{T} \hat{\varepsilon}_t^2, \tag{7}$$

where l stands for the number of estimated coefficients and

$$\hat{\varepsilon}_t = y_t - \hat{\mu}^{RF} - \hat{\beta}^{RF} t - \hat{\alpha}_1^{RF} y_{t-1} - \cdots - \hat{\alpha}_p^{RF} y_{t-p}. \tag{8}$$

The point forecasts $\hat{y}(h)$ for $1, \ldots, H$ and the corresponding prediction standard errors $\hat{\sigma}(h)$ may be calculated according to

$$\hat{y}(h) = \hat{\mu}^{RF} + \hat{\beta}^{RF}(T+h) + \hat{\alpha}_1^{RF}\hat{y}(h-1) + \cdots + \hat{\alpha}_p^{RF}\hat{y}(h-p), \tag{9}$$

where $\hat{y}(j) = y_{T+j}$ for $j \leq 0$,

and

$$\hat{\sigma}(h) = \hat{\sigma}^2 \sqrt{\hat{\theta}_0^2 + \cdots + \hat{\theta}_{h-1}^2}, \tag{10}$$

where $\hat{\theta}_j = \sum_{i=1}^{j} \hat{\theta}_{j-i}\, \hat{\alpha}_i^{RF}$ for $j = 1,2, \ldots, h-1$ with $\hat{\alpha}_i^{RF} = 0$ for $i > p$ and $\hat{\theta}_0 = 1$.

In what follows the main interest lies in using the Roy-Fuller estimator and the bootstrap method for constructing prediction bands which should cover the future *H*-dimensional path of realizations $y(H) = (y_{T+1}, \ldots, y_{T+H})'$ which a preassigned probability.

3. THE BOOTSTRAP ALGORITHM

Predictive distributions and estimates of standardized prediction errors are obtained using the residual bootstrap procedure. Calculations involve a number of steps:

1) The parameters of model (1) are estimated using the Roy-Fuller method and the corresponding residuals are computed. The residuals from (8) are inflated using a factor of $\sqrt{\frac{T}{T-l}}$ (see e.g. Stine, 1987) and denoted by $\hat{\varepsilon}_t^*$.
2) A sample of pseudo-data of size T is generated from the bootstrap data generating process of the form (see Clements, Kim, 2007; Fresoli et al., 2015):

$$y_t^* = \hat{\mu}^{RF} + \hat{\beta}^{RF} t + \hat{\alpha}_1^{RF} y_{t-1}^* + \cdots + \hat{\alpha}_p^{RF} y_{t-p}^* + \varepsilon_t^*, \tag{11}$$

where actual observations $y_{-p+1}, \ldots, y_0$ are used as pre-sample values $y_{-p+1}^*, \ldots, y_0^*$ and ε_t^* is drawn randomly from the rescaled residual series $\hat{\varepsilon}_t^*$.

3) The pseudo-data set is used to re-estimate the parameters of model (1) producing $\hat{\mu}^{RF*}, \hat{\beta}^{RF*}, \hat{\alpha}_1^{RF*}, \ldots, \hat{\alpha}_p^{RF*}$ and also to compute forecasts $\hat{y}^*(h)$ and prediction standard errors $\hat{\sigma}^*(h)$ as in (9) and (10) but with $\hat{\mu}^{RF}, \hat{\beta}^{RF}, \hat{\alpha}_1^{RF}, \ldots, \hat{\alpha}_p^{RF}$ replaced by $\hat{\mu}^{RF*}, \hat{\beta}^{RF*}, \hat{\alpha}_1^{RF*}, \ldots, \hat{\alpha}_p^{RF*}$.
4) Bootstrap future trajectory for horizon H, $(y^*(1), \ldots, y^*(H))'$ is generated using:

$$\begin{aligned} y^*(h) = \hat{\mu}^{RF*} + \hat{\beta}^{RF*}(T+h) + \hat{\alpha}_1^{RF*} y^*(h-1) + \cdots + \\ + \hat{\alpha}_p^{RF*} y^*(h-p) + \varepsilon_h^*, \end{aligned} \tag{12}$$

where $h = 1, \ldots, H$, $y^*(i) = y_{T+i}$ for $i \leq 0$ and ε_h^* is drawn randomly from the series $\hat{\varepsilon}_t^*$.

5) Bootstrap vector of standardized prediction errors $\hat{S}^*(H) = (\hat{s}^*(1), \ldots, \hat{s}^*(H))'$ is evaluated by generating y_t^* for $T+1 \leq t \leq T+H$ analogously as in (11) but with initial values given by $y_{T-p+1}, \ldots, y_T$ and calculating, for $h = 1, \ldots, H$:

$$\hat{s}^*(h) = \frac{\hat{y}^*(h) - y^*(h)}{\hat{\sigma}^*(h)}. \tag{13}$$

The procedure in steps (2)–(5) is repeated *N* times (where *N* denotes the number of iterations in the bootstrap loop), providing *N* bootstrap replicates of the future trajectory of y and the same number of replicates of the vector of standardized prediction errors. These values can be used to construct various bootstrap prediction bands.

4. REFINED BONFERRONI BANDS

Prediction bands can be constructed using Bonferroni's method. Suppose the objective is to construct a $(1-\gamma)\times 100\%$ prediction band *B* for the elements of an *H*-dimensional vector $y(H)=(y_{T+1},\dots,y_{T+H})'$. The Bonferroni inequality

$$P(y(H)\in B)\geq \sum_{h=1}^{H}(1-\gamma_h)-(H-1), \tag{14}$$

where $1-\gamma_h$ is the coverage probability of the band at horizon h (i.e. computed with respect to y_{T+h}), indicates that in order to achieve at least the desired coverage of the band, it can be assumed that $\sum_{h=1}^{H}\gamma_h=\gamma$. The most common approach is to set each γ_h to the same value $\frac{\gamma}{H}$. The resulting band is constructed from $\left(1-\frac{\gamma}{H}\right)\times 100\%$ prediction intervals for each element y_{T+h} separately:

$$B=[B_{L1},B_{U1}]\times\ [B_{L2},B_{U2}]\times\dots\times[B_{LH},B_{UH}], \tag{15}$$

where B_{Lh} and B_{Uh} denote respectively, the $\frac{\gamma}{2H}$ and $1-\frac{\gamma}{2H}$ quantiles of the predictive distribution of y_{T+h}.

Given that the actual coverage of the Bonferroni band may easily exceed the desired level and the band may, in effect, be excessively wide (see, e.g. Lütkepohl et al., 2015*b*) it makes sense to try to refine the Bonferroni bands in such a way that the actual coverage becomes closer to the nominal level and the width of the bands is reduced.

The first refinement uses higher order Bonferroni-type inequalities of Hoover (1990). Glaz, Ravishanker (1991) apply these inequalities to construct prediction bands for ARIMA models using the properties of the multivariate normal distribution. In this paper, the bootstrap distribution of the predictor is considered. The condition implied by the Bonferroni-type inequality of order *k*, for $1<k\leq H-1$ for the band B^k has the form:

$$P(y(H)\in B^k)\geq \sum_{h=1}^{H+1-k}\left(1-\gamma_{h,h+k-1}\right)-\sum_{h=1}^{H-k}\left(1-\gamma_{h+1,h+k-1}\right), \tag{16}$$

where for $1\leq m\leq n\leq H$, $(1-\gamma_{m,n})$ is the probability that realizations of the predicted variable observed from horizon *m* to horizon *n* is covered by the corresponding stretch of the band, $\gamma_{m,m}=\gamma_m$ and $\gamma_{m+1,m}\equiv 1$.

The band with equal values of γ_h for $h=1,\dots,H$, could be constructed in an iterative manner by starting from the band derived from Bonferroni's inequality

and reducing the coverage of the intervals in each step by 1/*N*, where *N* is the number of bootstrap replications (making γ_h larger in each step by 1/*N*). In each iteration it would be checked whether the relevant higher order Bonferroni inequality was observed. This can be achieved by computing the bootstrap coverage at single horizons and multiple horizons and evaluating the Bonferroni-type inequality of order k of interest. The final values of γ_h would be the largest values for which the inequality was met. The larger the value of k, the less conservative the resulting band.

The second refinement of the basic Bonferroni band aims at finding such γ_h for $h = 1, \dots, H$, $\sum_{h=1}^{H} \gamma_h = \gamma$ that the width of the band is as small as possible. The resulting Bonferroni band can be described as imbalanced (see e.g. Wolf, Wunderli, 2015).

Optimization is done using threshold accepting (TA) belonging to a class of refined local search methods. The procedure was proposed by Dueck, Scheuer (1990) and applied to the problem of constructing prediction bands e.g. by Staszewska-Bystrova, Winker (2013) and Grabowski et al. (2017).

The objective function which is minimized is the width of the band:

$$W(B^{TA}) = \sum_{h=1}^{H} (B_{Uh}^{TA} - B_{Lh}^{TA}), \tag{17}$$

where B^{TA} is the Bonferroni band obtained using threshold accepting. The detailed steps of the optimization procedure are presented in algorithm 1.

The algorithm is initialized (step 1) by considering as the starting solution B^c, the basic Bonferroni band *B* and evaluating the objective function for this band. The number of search steps (n_{iter}) and the threshold sequence ($t_1, \dots, t_{n_{iter}}$) of the corresponding length are also set. The threshold values should be positive and decreasing. Then n_{iter} iterations are performed. In each iteration i, a new solution (prediction band) belonging to the neighborhood of the current solution is considered. The neighboring band is created randomly by modifying the width of the currently considered band in two points in such a way that the constraint $\sum_{h=1}^{H} \gamma_h = \gamma$ is not violated. To achieve this, two values h_1 and h_2 are randomly selected from the set $\{1, \dots, H\}$ (step 2) and the corresponding values γ_{h_1} and γ_{h_2} are changed (steps 3–5). First, γ_{h_1} is made smaller by subtracting a random fraction of its current value (f). Second, γ_{h_2} is enlarged by f. Then, the width of the new band B^n is computed and compared to the width of the current band (step 6). If the difference is smaller than the threshold value for iteration i, then the solution is accepted as the current solution (step 7). The algorithm continues until n_{iter} search steps are completed. The band with the smallest value of the objective function found throughout the search steps, denoted by B^{TA} is presented as the final solution.

Algorithm 1.

Threshold accepting procedure

1. Obtain initial solution B^c and compute $W(B^c)$. Set the value of n_{iter} and $t_1, \dots, t_{n_{iter}}$
2. for $i = 1, \dots, n_{iter}$ do
3. Randomly select two integers: h_1 and h_2 from the set $\{1, \dots, H\}$
4. Randomly select *p* from the interval (0,1) and compute $f = p\gamma_{h_1}^c$
5. Obtain new solution B^n by setting $\gamma_{h_1}^n = \gamma_{h_1}^c - f$ and $\gamma_{h_2}^n = \gamma_{h_2}^c + f$
6. Calculate $\Delta = W(B^n) - W(B^c)$
7. if $\Delta < t_i$ then $B^c = B^n$
8. end for

The proposed Bonferroni-type bands are compared to the sup-*t* method described by Wolf, Wunderli (2015). The benchmark procedure has some optimality properties (e.g smaller width) as compared to the traditional Bonferroni algorithm in large samples (see Montiel Olea, Plagborg-Møller, 2017). The sup-*t* bands are computed by finding the largest value in each of the N vectors $|\hat{S}^*(H)|$ and obtaining $d_{1-\alpha}$ equal to the $1-\alpha$ quantile of these maxima. In the next step the band is formed as

$$[\hat{y}(1) \pm d_{1-\alpha}\hat{\sigma}(1)] \times \dots \times [\hat{y}(H) \pm d_{1-\alpha}\hat{\sigma}(H)]. \tag{18}$$

5. A SIMULATION STUDY

Small-sample properties of the bands were studied using Monte Carlo simulations. A number of data generating processes (DGPs) were investigated. The first set of DGPs (denoted by DGP.A), considered also by Clements, Kim (2007), had the form:

$$y_t = 1 + (1-\alpha)t + \alpha y_{t-1} + \varepsilon_t, \tag{19}$$

where $\alpha \in \{0.5, 0.9, 0.95\}$. Higher values of α correspond to larger degree of persistence of the AR process. Three different distributions of the errors, ε_t, were considered for DGP.A: a standard normal distribution ($N(0,1)$), a chi-square distribution with 4 degrees of freedom, centered to have mean 0 and standardized to have variance 1 and a *t*-distribution with 4 degrees of freedom standardized to have variance equal to 1.

More complex DGPs (DGP.B), corresponding to AR(2) were given by:

$$y_t = 1 + (1.85 - \alpha_1)t + \alpha_1 y_{t-1} - 0.85 y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim N(0,1), \tag{20}$$

where $\alpha_1 \in \{1.35, 1.75, 1.8\}$. For these processes $\gamma_1 = \alpha_1 - 0.85$ and so the persistence also grows as the value of α_1 increases.

The other settings of the Monte Carlo experiments were as follows. The number of Monte Carlo replications *M* was set to 1000 and the number of iterations in the bootstrap procedure *N* was equal to 2000.[4] In each Monte Carlo iteration, parameters of an AR model with a constant and trend were estimated. The lag order was selected using Akaike's information criterion (AIC) allowing for up to 8 lags. The same number of lags was used for the models estimated in the bootstrap procedure. The sample size *T* and the forecast horizon *H* belonged respectively to the following sets: $T \in \{100, 400\}$ and $H \in \{4, 8, 12\}$. The nominal coverage rate of the bands was given by 0.9.

Further parameter settings corresponded to specific methods of constructing bands. The value of k for the procedure based on higher order Bonferroni's inequality was equal to 2, 3 or 4. Since the combination $k = 4$ and $H = 4$ is not feasible due to the condition $k \leq H - 1$, the results were not obtained for these cases and the corresponding entries in tables are given as NA. TA optimization was performed for $n_{iter} = 500000$ and the threshold sequence defined as $t_i = \frac{n_{iter}-i}{n_{iter}} \times 0.05$ for $i = 1, \ldots, n_{inter}$.

Two properties of prediction bands were evaluated in the simulations: mean coverage rates and average width. In order to evaluate the coverage probabilities, 1000 future trajectories of length *H* (each computed conditionally on the last *p* values from the generated sample) were obtained from the DGPs. Then, in every Monte Carlo replication the proportions of trajectories lying entirely within the alternative prediction bands were computed. Mean coverage rates were obtained as averages of these proportions over *M* replications. To provide a measure of width of the bands, sum of differences between the upper and lower bounds were calculated for $h = 1, \ldots, H$ and divided by *H*. Average values for *M* Monte Carlo iterations are reported.

The results of all experiments are presented in tables 1–6 and tables A1–A6 from the Appendix. Tables 1–3 and 4–6 contain results obtained respectively for DGP.A with normal errors and DGP.B with alternative parameter values. In tables A1–A3 and A4–A6 estimated coverage rates and width measures are presented for DGP.A with chi-square distributed and *t*-distributed errors. Quantities without parentheses correspond to the estimated coverage probabilities, while values in parentheses indicate average width of the bands.

Table 1. RESULTS FOR DGP.A WITH $\alpha = 0.5$ AND NORMAL ERRORS

H	*B*	sup-*t*	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 100}$			
4	89.16 (5.18)	88.55 (5.07)	87.83 (5.04)	87.59 (5.02)	NA	88.58 (5.12)

[4] Some experiments showed that using 5000 Monte Carlo and 5000 bootstrap iterations did not change the conclusions significantly.

Table 1. RESULTS FOR DGP.A WITH $\alpha = 0.5$ AND NORMAL ERRORS (cont.)

H	*B*	sup-*t*	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 100}$ **(dok.)**			
8	88.67 (6.01)	88.27 (5.87)	86.93 (5.82)	86.68 (5.80)	86.55 (5.79)	87.90 (5.93)
12	88.06 (6.45)	88.09 (6.31)	86.22 (6.26)	85.92 (6.23)	85.79 (6.22)	87.20 (6.36)
			$\boldsymbol{T = 400}$			
4	90.51 (5.00)	89.68 (4.89)	89.35 (4.89)	89.15 (4.87)	NA	89.91 (4.95)
8	90.40 (5.72)	89.68 (5.61)	88.86 (5.58)	88.68 (5.56)	88.59 (5.55)	89.53 (5.66)
12	90.07 (6.10)	89.59 (5.99)	88.42 (5.95)	88.29 (5.94)	88.21 (5.93)	89.00 (6.02)

The Monte Carlo results for the proposed versions of the Bonferroni band can be summarized as follows. While all the refinements work in the expected way and bring down the average width of the prediction bands as compared to the Bonferroni band, the size of the reduction differs between methods and depends on the specific features of the DGP and the sample size. It has also varying impact on the estimated coverage probabilities.

A general observation is that B^{TA} is almost always wider than the widest of the B^k bands, i.e. B^2. The three versions of the bands based on higher order Bonferroni-type inequalities do not differ much in terms of width. Given that B^2 is considerably more aggressive than the Bonferroni band and that further small reductions in width as implied by B^3 and B^4 tend to impair the coverage probabilities of these bands, B^2 might be preferred over the remaining B^k methods. The length of the forecast horizon has an expected impact on the width of the prediction bands for all the methods, i.e. the width grows as H increases, however it does not influence the relative ordering of the Bonferroni-type procedures for band construction.

Table 2. RESULTS FOR DGP.A WITH $\alpha = 0.9$ AND NORMAL ERRORS

H	*B*	sup-*t*	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 100}$			
4	91.36 (6.93)	89.03 (6.46)	88.57 (6.43)	88.15 (6.37)	NA	90.91 (6.80)
8	91.77 (9.90)	88.59 (8.96)	88.13 (8.94)	87.46 (8.81)	87.18 (8.76)	91.46 (9.69)
12	91.74 (12.16)	88.08 (10.85)	87.84 (10.84)	12.20 (10.66)	86.72 (10.58)	91.51 (11.90)

Table 2. RESULTS FOR DGP.A WITH $\alpha = 0.9$ AND NORMAL ERRORS (cont.)

H	*B*	sup-*t*	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 400}$			
4	92.32 (6.43)	89.77 (6.02)	89.71 (6.05)	89.34 (6.00)	NA	91.88 (6.34)
8	93.28 (8.71)	89.81 (7.97)	89.92 (8.06)	89.30 (7.97)	89.07 (7.93)	92.88 (8.57)
12	93.66 (10.20)	89.81 (9.21)	90.08 (9.40)	89.41 (9.28)	89.10 (9.23)	93.26 (10.03)

Table 3. RESULTS FOR DGP.A WITH $\alpha = 0.95$ AND NORMAL ERRORS

H	*B*	sup-*t*	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 100}$			
4	91.54 (7.14)	89.04 (6.62)	88.62 (6.59)	88.18 (6.52)	NA	91.05 (7.00)
8	92.06 (10.57)	88.47 (9.47)	88.19 (9.42)	87.48 (9.28)	87.19 (9.22)	91.72 (10.33)
12	91.99 (13.32)	87.70 (11.76)	87.82 (11.69)	87.00 (11.48)	86.61 (11.39)	91.73 (13.01)
			$\boldsymbol{T = 400}$			
4	92.63 (6.73)	89.81 (6.24)	89.85 (6.29)	89.44 (6.23)	NA	92.24 (6.63)
8	93.84 (9.59)	89.90 (8.62)	90.21 (8.78)	89.58 (8.66)	89.30 (8.62)	93.55 (9.42)
12	94.41 (11.73)	89.93 (10.33)	90.55 (10.64)	89.81 (10.48)	89.47 (10.42)	94.16 (11.51)

A more detailed assessment of all the methods leads to the following observations. Results obtained for DGP.A with normal errors (tables 1–3) indicate that performance of the procedures depends on the persistence of the process that generated the data and the available number of observations. These features influence the working of the basic Bonferroni band which becomes more conservative for more persistent processes and larger samples which in turn brings about the need for refinement and also the sup-*t* method which tends to undercover for the smaller sample size, especially if the forecast horizon is long. For $\alpha = 0.5$ (table 1) the B band is overall best for both sample sizes in terms of the estimated coverage probabilities, however for bigger values of α (tables 2–3) these probabilities become too large as compared to the nominal rate of 0.9. This effect can be observed for both smaller sample size and larger sample size, where as expected, it is even more pronounced. In these cases, selected alter-

native methods may be preferable. In particular, if the number of observations is small, B^{TA} bands could be used as they maintain the nominal coverage rate, while for larger sample sizes B^2 and sup-*t* bands have the best coverage properties.[5]

Table 4. RESULTS FOR DGP.B WITH $\alpha_1 = 1.35$

H	*B*	sup-*t*	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 100}$			
4	90.41 (7.68)	88.62 (7.29)	87.27 (7.15)	87.19 (7.14)	NA	89.78 (7.54)
8	89.55 (10.00)	88.55 (9.72)	86.44 (9.44)	86.41 (9.44)	86.11 (9.39)	88.72 (9.82)
12	89.12 (11.58)	88.57 (11.34)	86.23 (11.00)	86.19 (11.00)	85.77 (10.93)	88.28 (11.37)
			$\boldsymbol{T = 400}$			
4	92.13 (7.45)	89.69 (6.99)	89.18 (6.95)	89.10 (6.94)	NA	91.63 (7.33)
8	92.16 (9.83)	89.74 (9.30)	89.37 (9.29)	89.34 (9.29)	89.00 (9.23)	91.41 (9.67)
12	92.10 (11.39)	89.74 (10.80)	89.50 (10.84)	89.47 (10.83)	89.10 (10.76)	91.30 (11.20)

Table 5. RESULTS FOR DGP.B WITH $\alpha_1 = 1.75$

H	*B*	sup-*t*	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 100}$			
4	92.39 (11.55)	88.85 (10.35)	87.61 (10.18)	87.42 (10.14)	NA	91.50 (11.17)
8	92.53 (20.13)	88.21 (17.63)	86.09 (17.23)	85.87 (17.15)	85.78 (17.13)	91.98 (19.51)
12	92.20 (25.08)	87.94 (21.92)	85.28 (21.44)	85.07 (21.36)	84.99 (21.33)	91.67 (24.41)
			$\boldsymbol{T = 400}$			
4	94.02 (11.15)	89.80 (9.85)	89.49 (9.83)	89.31 (9.79)	NA	93.45 (10.83)
8	94.85 (18.87)	89.70 (16.29)	89.09 (16.24)	88.90 (16.18)	88.83 (16.16)	94.49 (18.37)
12	94.82 (22.90)	89.62 (19.91)	88.85 (19.87)	88.70 (19.81)	88.64 (19.79)	94.35 (22.38)

[5] Some additional simulation results obtained for DGP.A with $\alpha = 0.9$ indicate that B^{TA} bands might be considered as superior to B^2 and sup-*t* bands for sample sizes smaller than 130, where the mean coverage of the latter bands falls below 0.89 for longer forecast horizons.

Table 6. RESULTS FOR DGP.B WITH $\alpha_1 = 1.8$

H	*B*	sup-*t*	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 100}$			
4	92.63 (12.28)	88.82 (10.89)	87.65 (10.73)	87.46 (10.69)	NA	91.64 (11.84)
8	93.19 (24.11)	87.97 (20.55)	86.26 (20.19)	85.97 (20.07)	85.85 (20.03)	92.54 (23.23)
12	92.96 (33.44)	87.22 (28.06)	85.04 (27.58)	84.76 (27.44)	84.65 (27.38)	92.43 (32.35)
			$\boldsymbol{T = 400}$			
4	94.19 (11.80)	89.81 (10.35)	89.52 (10.35)	89.33 (10.30)	NA	93.56 (11.43)
8	95.34 (22.44)	89.75 (18.94)	89.26 (18.96)	89.01 (18.86)	88.92 (18.83)	95.03 (21.74)
12	95.60 (29.94)	89.66 (25.06)	89.03 (25.13)	88.80 (25.02)	88.71 (24.97)	95.29 (29.11)

Conclusions from the results for DGP.B (tables 3–6) are similar as those for DGP.A. Refined and the sup-*t* methods are most useful in larger samples and for predicting persistent processes for which the Bonferroni band tends to be too wide and have excessive coverage probability. The largest gains can be obtained for larger samples for the sup-*t* and B^2 methods, for which the reduction in band width can be considerable (e.g. more than 15% for $\alpha_1 = 1.8, T = 400$ and $H = 12$). At the same time the coverage probabilities for these procedures are quite close to 0.9, especially in the case of the sup-*t* method. For prediction based on persistent processes and smaller data sets B^{TA} method could be selected.

As follows from the analysis of tables A1–A6, the findings for normal DGPs are to some extent, robust with respect to the distribution of the random errors. For more persistent processes with either chi-square errors or *t*-distributed errors considered in tables A2–A3 and A5–A6, as previously, the B^{TA} procedure could be considered as most robust for all values of *H* for the smaller sample size, while sup-*t* or B^2 bands, which have very similar properties, would be the natural choice for larger samples. Some new effects can be observed for DGP.A with $\alpha = 0.5$, however. For the process with chi-square innovations (table A1), the Bonferroni bands still have the best coverage properties for $T = 100$, however there is a new winner for $T = 400$ given by the B^{TA} procedure. The B^{TA} bands do not undercover for any value of *H* and are narrower as compared to the Bonferroni bands (by construction) and to the sup-*t* bands. As table A4 reveals, a different method, namely the sup-*t* procedure is most reliable for this less persistent process with fat-tailed distribution of the error terms for both $T = 100$ and $T = 400$.

6. CONCLUSIONS

Joint prediction bands are needed for forming expectations concerning the future trajectory of a variable. The construction of such bands is usually based on the bootstrap predictive distribution. One of the classic approaches to build prediction bands rests on the Bonferroni inequality. The drawback of this method is that the bands can be too wide and exhibit larger probability content than the nominal coverage rate.

In this study two refinements of the Bonferroni band were considered in the context of predicting persistent univariate autoregressive processes. The first refinement used higher order Bonferroni-type inequalities, while the second consisted in constructing the band from intervals with unequal coverage rates. The proposed bands were compared to the Bonferroni bands and the benchmark given by the sup-*t* procedure in a Monte Carlo study.

Simulation results indicated that the refined bands were superior to the basic Bonferroni bands in a number of scenarios involving quite persistent processes. In particular, bands based on the second-order Bonferroni-type inequality worked well for relatively large samples where they exhibited similar properties as the sup-*t* bands, while the imbalanced Bonferroni method was preferable for smaller sample sizes.

REFERENCES

Andrews D. W. K., Chen H.-Y., (1994), Approximate Median-Unbiased Estimation of Autoregressive Models, *Journal of Business & Economic Statistics*, 12, 187–204.

Box G. E. P., Jenkins G. M., (1970), *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco.

Breidt F. J., Davis R. A., Dunsmuir W., (1995), Improved Bootstrap Prediction Intervals for Autoregressions, *Journal of Time Series Analysis*, 16, 177–200.

Clements M. P., Kim J. H., (2007), Bootstrap Prediction Intervals for Autoregressive Time Series, *Computational Statistics & Data Analysis*, 51, 3580-3594.

Dueck, G., Scheuer T., (1990), Threshold Accepting: A General Purpose Algorithm Appearing Superior to Simulated Annealing, *Journal of Computational Physics*, 90, 161–175.

Efron B., (1979), Bootstrap Methods: Another Look at the Jackknife, *The Annals of Statistics*, 7, 1–26.

Fresoli D., Ruiz E., Pascual L., (2015), Bootstrap Multi-Step Forecasts of Non-Gaussian VAR Models, *International Journal of Forecasting*, 31, 834–848.

Glaz J., Ravishanker N., (1991), Simultaneous Prediction Intervals for Multiple Forecasts Based on Bonferroni and Product-Type Inequalities, *Statistics & Probability Letters*, 12, 57–63.

Grabowski D. Staszewska-Bystrova A., Winker P., (2017), Generating Prediction Bands for Path Forecasts from SETAR Models, *Studies in Nonlinear Dynamics & Econometrics*, 21, 5.

Grigoletto M., (1998), Bootstrap Prediction Intervals for Autoregressions: Some Alternatives, *International Journal of Forecasting*, 14, 447–456.

Hoover D. R., (1990), Subset Complement Addition Upper Bounds – An Improved Inclusion-Exclusion Method, *Journal of Statistical Planning and Inference*, 24, 195–202.

Jordà Ò., Marcellino M., (2010), Path Forecast Evaluation, *Journal of Applied Econometrics*, 25, 635–662.

Kilian L., (1998), Small-Sample Confidence Intervals for Impulse Response Functions, *Review of Economics and Statistics*, 80, 218–230.

Kim J. H., (2001), Bootstrap-after-Bootstrap Prediction Intervals for Autoregressive Models, *Journal of Business & Economic Statistics*, 19, 117–128.

Lütkepohl H., Krätzig M., (2004), *Applied Time Series Econometrics*, Cambridge University Press, Cambridge.

Lütkepohl H., Staszewska-Bystrova A., Winker P., (2015*a*), Comparison of Methods for Constructing Joint Confidence Bands for Impulse Response Functions, *International Journal of Forecasting*, 31, 782–798.

Lütkepohl H., Staszewska-Bystrova A., Winker P., (2015*b*), Confidence Bands for Impulse Responses: Bonferroni versus Wald, *Oxford Bulletin of Economics and Statistics*, 77, 800–821.

Masarotto G. (1990), Bootstrap Prediction Intervals for Autoregressions, *International Journal of Forecasting* 6, 229–239.

Montiel Olea J. L., Plagborg-Møller M., (2017), Simultaneous Confidence Bands: Theoretical Comparisons and Suggestions for Practice, *Working Paper*.

Roy A., Fuller W. A., (2001), Estimation for Autoregressive Time Series with a Root Near One, *Journal of Business & Economic Statistics*, 19, 482–493.

Staszewska-Bystrova A., (2011), Bootstrap Prediction Bands for Forecast Paths from Vector Autoregressive Models, *Journal of Forecasting*, 30, 721–735.

Staszewska-Bystrova A., (2013), Modified Scheffè's Prediction Bands, *Jahrbücher für Nationalökonomie und Statistik*, 233, 680–690.

Staszewska-Bystrova A., Winker P., (2013), Constructing Narrowest Pathwise Bootstrap Prediction Bands using Threshold Accepting, *International Journal of Forecasting*, 29, 221–233.

Stine R. A., (1987), Estimating Properties of Autoregressive Forecasts, *Journal of the American Statistical Association,* 82, 1072–1078.

Thombs L. A., Schucany W. R., (1990), Bootstrap Prediction Intervals for Autoregression, *Journal of the American Statistical Association*, 85,. 486–492.

Wolf M., Wunderli D., (2015), Bootstrap Joint Prediction Regions, *Journal of Time Series Analysis*, 36, 352–376.

APPENDIX

A1.

Table RESULTS FOR DGP.A WITH $\alpha = 0.5$ AND X^2-DISTRIBUTED ERRORS

H	*B*	sup-*t*	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 100}$			
4	90.54 (5.28)	88.85 (5.34)	89.16 (5.08)	88.93 (5.05)	NA	89.83 (5.18)
8	90.09 (6.23)	88.07 (6.63)	88.43 (6.01)	88.18 (5.98)	88.05 (5.96)	88.99 (6.10)
12	89.14 (6.67)	87.45 (7.40)	87.47 (6.46)	87.22 (6.43)	87.09 (6.41)	88.09 (6.56)

A1. (cont.)

Table RESULTS FOR DGP.A WITH $\alpha = 0.5$ AND X^2-DISTRIBUTED ERRORS (cont.)

H	*B*	sup-*t*	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 400}$			
4	91.45 (5.01)	89.73 (5.15)	90.19 (4.86)	89.97 (4.84)	NA	90.86 (4.93)
8	90.06 (6.82)	89.78 (6.48)	88.47 (6.43)	88.35 (6.40)	88.28 (6.39)	90.67 (5.75)
12	91.40 (6.32)	89.43 (7.14)	89.65 (6.10)	89.51 (6.08)	89.41 (6.07)	90.32 (6.18)

A2.

Table RESULTS FOR DGP.A WITH $\alpha = 0.9$ AND X^2-DISTRIBUTED ERRORS

H	*B*	sup-*t*	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 100}$			
4	92.01 (6.97)	89.43 (6.55)	89.25 (6.39)	88.84 (6.32)	NA	91.37 (6.83)
8	92.14 (10.02)	88.98 (9.34)	88.65 (8.95)	88.01 (8.81)	87.74 (8.75)	91.60 (9.79)
12	91.87 (12.27)	88.57 (11.37)	88.18 (10.87)	87.49 (10.68)	87.15 (10.59)	91.44 (12.01)
			$\boldsymbol{T = 400}$			
4	92.92 (6.42)	89.95 (5.94)	90.26 (5.98)	89.87 (5.92)	NA	92.45 (6.30)
8	93.89 (8.78)	89.93 (8.15)	90.48 (8.05)	89.92 (7.95)	89.68 (7.91)	93.39 (8.62)
12	94.17 (10.30)	89.92 (9.58)	90.59 (9.41)	89.95 (9.29)	89.65 (9.24)	93.66 (10.11)

A3.

Table RESULTS FOR DGP.A WITH $\alpha = 0.95$ AND X^2-DISTRIBUTED ERRORS

H	*B*	sup-*t*	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 100}$			
4	92.25 (7.15)	89.66 (6.66)	89.38 (6.52)	88.96 (6.45)	NA	91.63 (6.99)
8	92.46 (10.62)	89.03 (9.76)	88.69 (9.38)	88.04 (9.23)	87.75 (9.17)	91.94 (10.36)
12	92.17 (13.31)	88.24 (12.15)	88.16 (11.64)	87.38 (11.42)	87.03 (11.32)	91.78 (13.00)
			$\boldsymbol{T = 400}$			
4	93.25 (6.71)	89.98 (6.13)	90.37 (6.21)	89.95 (6.15)	NA	92.77 (6.59)
8	94.46 (9.66)	89.99 (8.75)	90.79 (8.74)	90.15 (8.63)	89.86 (8.57)	94.06 (9.47)
12	94.91 (11.81)	90.01 (10.62)	90.98 (10.62)	90.24 (10.45)	89.91 (10.38)	94.52 (11.58)

A4.

Table RESULTS FOR DGP.A WITH $\alpha = 0.5$ AND *t*-DISTRIBUTED ERRORS

H	*B*	sup-*t*	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 100}$			
4	88.69 (5.71)	88.87 (5.43)	87.51 (5.44)	87.31 (5.40)	NA	88.04 (5.56)
8	86.77 (7.03)	88.00 (6.76)	85.58 (6.82)	85.43 (6.79)	85.34 (6.78)	85.58 (6.78)
12	84.40 (7.52)	87.16 (7.68)	83.13 (7.33)	82.97 (7.31)	82.88 (7.29)	83.32 (7.34)
			$\boldsymbol{T = 400}$			
4	90.68 (5.47)	90.03 (5.25)	89.38 (5.22)	89.21 (5.19)	NA	90.12 (5.37)
8	90.06 (6.82)	89.78 (6.48)	88.47 (6.43)	88.35 (6.40)	88.28 (6.39)	89.25 (6.63)
12	89.36 (7.71)	89.56 (7.26)	87.85 (7.25)	87.71 (7.22)	87.64 (7.20)	88.35 (7.41)

A5.

Table RESULTS FOR DGP.A WITH $\alpha = 0.9$ AND *t*-DISTRIBUTED ERRORS

H	*B*	sup-*t*	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 100}$			
4	91.29 (7.39)	89.62 (6.62)	88.42 (6.60)	88.06 (6.53)	NA	90.67 (7.20)
8	91.31 (10.43)	89.14 (9.35)	87.71 (9.23)	87.13 (9.09)	86.89 (9.03)	90.72 (10.16)
12	90.81 (12.58)	88.56 (11.35)	87.08 (11.10)	86.39 (10.92)	86.09 (10.83)	90.42 (12.29)
			$\boldsymbol{T = 400}$			
4	92.93 (6.97)	90.10 (6.12)	89.74 (6.14)	89.42 (6.07)	NA	92.53 (6.83)
8	93.67 (9.93)	90.19 (8.35)	89.82 (8.48)	89.32 (8.35)	89.11 (8.30)	93.10 (9.65)
12	93.73 (11.64)	90.15 (9.80)	89.93 (10.06)	89.36 (9.89)	89.10 (9.82)	93.11 (11.35)

A6.

Table RESULTS FOR DGP.A WITH $\alpha = 0.95$ AND *t*-DISTRIBUTED ERRORS

H	*B*	sup-*t*	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 100}$			
4	91.61 (7.62)	89.77 (6.77)	88.62 (6.75)	88.24 (6.67)	NA	90.98 (7.42)
8	91.75 (11.10)	89.21 (9.86)	87.86 (9.68)	87.26 (9.52)	87.01 (9.46)	91.19 (10.80)
12	91.30 (13.71)	88.41 (12.25)	87.24 (11.91)	86.46 (11.68)	86.11 (11.58)	90.90 (13.38)

A6. (cont.)

Table RESULTS FOR DGP.A WITH $\alpha = 0.95$ AND t-DISTRIBUTED ERRORS (cont.)

H	B	sup-t	B^2	B^3	B^4	B^{TA}
			$\boldsymbol{T = 400}$			
4	93.25 (7.29)	90.11 (6.32)	89.81 (6.35)	89.47 (6.28)	NA	92.88 (7.14)
8	94.26 (10.86)	90.21 (8.94)	90.04 (9.12)	89.48 (8.97)	89.23 (8.90)	93.75 (10.56)
12	94.56 (13.19)	90.22 (10.83)	90.34 (11.19)	89.66 (10.97)	89.36 (10.88)	94.02 (12.86)

ZMODYFIKOWANE PASMA PREDYJKCYJNE BONFERRONIEGO DLA MODELI AUTOREGRESYJNYCH

Streszczenie

*Pasma predykcyjne konstruuje się często z użyciem nierówności Bonferroniego. Wadą takich pasm może być ich duża rozpiętość i zawyżone prawdopodobieństwo zawierania przyszłej trajektorii prognozowanej zmiennej. W artykule zaproponowano dwie poprawki dla metody konstrukcji bootstrapowych pasm predykcyjnych Bonferroniego wykorzystujące nierówności wyższego rzędu i procedurę minimalizacji szerokości pasma. Metody zastosowano do prognozowania jednowymiarowych procesów autoregresyjnych. Ich właściwości zbadano za pomocą eksperymentów Monte Carlo. Wykazano, że zaproponowane procedury prowadzą, w wielku przypadkach, do uzyskania stosunkowo wąskich pasm predykcyjnych o odpo*wiednich prawdopodobieństwach zawierania przyszłej trajektorii zmiennej.

Słowa kluczowe: pasmo predykcyjne, proces autoregresyjny, nierówność Bonferroniego

REFINED BONFERRONI PREDICTION BANDS FOR AUTOREGRESSIVE MODELS

Abstract

Joint prediction bands are often constructed using Bonferroni's inequality. The drawback of such bands may be their large width and excessive coverage probability. The paper proposes two refinements to the basic Bonferroni method of constructing bootstrap prediction bands. These are based on higher order inequalities and optimization of the width of the band. The procedures are applied to the problem of predicting univariate autoregressive processes. Their properties are studied by means of Monte Carlo experiments. It is shown that the proposed methods lead, in many scenarios, to obtaining relatively narrow prediction bands with desired coverage probabilities.

Keywords: prediction band, autoregressive process, Bonferroni's inequality