ISSN: 2278-5213

RESEARCH ARTICLE

Bootstrap control limits for Multivariate autocorrelated processes

A.A. Kalgonda
Dept. of Statistics, The New College, Kolhapur-416012, India
annagmk@rediffmail.com; +91 9922770979

Abstract

In recent years, due to automation of measurement and data collection systems, a process can be sampled at higher rates, which ultimately leads to autocorrelation. This has significant effect on the performance of classical control chart procedures. To handle multivariate process data, commonly, vector autoregressive model of order one (VAR(1) is used, which handle a wide range of autocorrelated process data. In this study, I propose evaluation of bootstrap control limits for vector autoregressive (VAR(1)) processes. This becomes useful in monitoring and diagnostic of multivariate autocorrelated processes. Example is given to evaluate performance of the suggested method.

Keywords: Automation, data collection, vector autoregressive model, multivariate autocorrelated process.

Introduction

In many instances of production processes, the quality of process depends upon more than one quality characteristic, which jointly determines the quality of the product. To deal with such processes, during recent years, an enormous amount of effort has been devoted to the study of multivariate control chart with interpretation of out of control signal when the data is serially independent. To name a few in multivariate control procedures I mention: Jackson (1985), Alt (1985) and Tracy *et al.* (1992). However, in practice, it has been observed that the assumption of independence is often violated and the data generated by the processes are autocorrelated. In such situation, classical control procedures perform poorly. A few multivariate statistical process control (SPC) methods have been developed when process observations are autocorrelated. Notably, Theodossiou (1993) proposed MCUSUM, Kramer and Schmid (1997) have proposed MEWMA. Stoumbos *et al*. (2000) have noted a great need of work on multivariate control charts with diagnosis when data is autocorrelated. In this direction, Kalgonda and Kulkarni (2004) have proposed Shewhart type chart (Z-chart) with diagnostic assuming that the underlying process follows VAR(1) model and error vector follows multivariate normal distribution. Control limits for the proposed test statistic are evaluated using critical value approach. The bootstrap approach is most appropriate tool in obtaining critical values of the test statistic (Efron and Tibshirani, 1998). Previously, bootstrap methods are used in univariate case by Seppala *et al.* (1995). Almost no work is reported in multivariate autocorrelated processes using bootstrap method.

Materials and methods

Evaluation of Bootstrap control limits

Assume that the underlying p-variate stationary process $\{\underline{X}_t\}$ exhibits a VAR (1) model, specified by,

$$\underline{X}_t = (I-\Phi)\underline{\mu} + \Phi\underline{X}_{t-1} + \in_t \quad t = 1,2,\ldots\ldots, \tag{1}$$

Where, $\underline{X}_t$ is the vector of process data at time t, $\underline{\mu}$ the vector of means and Φ is the autoregressive parameter matrix of order $p \times p$. Assume that the $\{\in_t\}$'s are independent and identically distributed multivariate normal, $N_p(\underline{0}, \Sigma)$. Then,

$$\underline{X}_t \sim N_p(\underline{\mu}, \Gamma(0))$$

Where $\Gamma(0) = ((\gamma_{ij}))$ is cross-covariance matrix.

When the process is in control, let the mean vector be $\underline{\mu}_0$ and the cross covariance matrix be $\Gamma(0)$. The Z-chart (Kalgonda and Kulkarni, 2004) is used for monitoring the mean of the VAR(1) process that consider the statistic

$$Z_t = \max\left\{\frac{\left|X_{it} - \mu_i^0\right|}{\gamma_i(0)}, \quad i = 1,2,\ldots,p\right\} \tag{2}$$

To obtain the control limits one need the critical value $C_{\rho(0),\alpha}$ that satisfies

$$\Pr\left\{Z_t \le C_{\rho(0),\alpha}\right\} = 1-\alpha \tag{3}$$

It is suggested to use simulation technique to evaluate $C_{\rho(0),\alpha}$. The bootstrap sampling procedure can be used effectively to obtain the critical values in such a situation.

Table 1. Estimated values of UCL using balanced bootstrap.

Sample	Type of UCL	Type I error				
		0.10	0.05	0.01	0.005	0.0026
A	$Z_B^*(\alpha)$	1.8588	2.1630	2.7727	2.988	3.2223
	$C_{\rho(0),\alpha}$	1.8760	2.1776	2.7781	3.0240	3.2312
	Bias	0.0172	0.0146	0.0054	0.036	0.0088
B	$Z_B^*(\alpha)$	1.8895	2.2105	2.7741	3.0107	3.3201
	$C_{\rho(0),\alpha}$	1.8755	2.1772	2.7912	3.0572	3.2573
	Bias	0.0196	0.0333	0.0171	0.0405	0.0628
C	$Z_B^*(\alpha)$	1.9326	2.2750	2.9988	3.2427	3.4230
	$C_{\rho(0),\alpha}$	1.9077	2.2064	2.7756	2.9689	3.1559
	Bias	0.0249	0.0686	0.2232	0.2738	0.2671

As there are several bootstrap tools available to evaluate critical value of test statistic for time series data, one appropriate tool is the 'balance bootstrap percentile' method (Davison and Hinkly, 1999). I use it here to obtain the critical values and hence the control limits as follows:

Step 1: Compute the residual vectors

$$\underline{e}_t = \underline{X}_t - (I-\Phi)\,\underline{\mu}_0 - \Phi\underline{X}_{t-1} \quad t = 2,3,\ldots\ldots,m \quad (4)$$

Step 2: Take B repetitions of residual vectors into a string of length, K-1

Such that,

$$K-1=(m-1)B, \quad (5)$$

generally, $K \geq 2000$ (Seppala *et al.,* 1995)

Step 3: Randomly permute the K-1 residual vectors denoting as bootstrap residuals $\underline{e}_t^*$, t=2,..,K.

Step 4: Let

$$\underline{X}_1^* = X_1$$

and (6)

$$\underline{X}_t^* = (I-\hat{\Phi})\hat{\underline{\mu}}^0 + \hat{\Phi}\underline{X}_{t-1}^* + \underline{e}_t^* \quad t = 2,3,\ldots,K \quad (7)$$

Step 5: Based on $\underline{X}_t^*$, compute the values of the statistic $Z_1^*, Z_2^*,\ldots\ldots,Z_K^*$

Where

$$Z_t^* = \max\left\{\left|\frac{X_{it}^* - \hat{\mu}_i^0}{\hat{\gamma}_i(0)}\right| \quad i=1,2,\ldots.,p\right\} \quad (8)$$

Step 6: Using K ordered observations of Z_t^* generated by B independent bootstrap residuals, take α^{th} quantile and obtain α level critical value denoting $Z_B^*(\alpha)$. Thus, $Z_B^*(\alpha)$ gives a value of $C_{\rho(0),\alpha}$, hence it can be used as UCL of the Z-chart.

Results and discussion

Example is given to evaluate performance of the suggested method.

Example

Consider bivariate normal time series data to assess the validity of bootstrap control limits with standard limits. I compare bootstrap control limits $Z_B^*(\alpha)$ with that obtained by simulation method $C_{\rho(0),\alpha}$ (Hayter and Tsui, 1994). For this purpose, I have considered three VAR(1) models with different parameters. Each time the observations were generated using the mean vector $\underline{\mu} = (0,0)'$ and the error covariance matrix $\Sigma = \begin{pmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{pmatrix}$.

The autocorrelated parameter matrices corresponding to the sample Sets A, B, and C were respectively taken as:

$$\Phi_A = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.7 \end{pmatrix}, \quad \Phi_B = \begin{pmatrix} 0.7 & 0 \\ 0 & 0.8 \end{pmatrix}, \quad \text{and} \quad \Phi_C = \begin{pmatrix} 0.7 & 0 \\ 0 & 0.9 \end{pmatrix}$$

The bootstrap percentile method is used to estimate the critical value. Each time 2000 bootstrap samples were generated and ordered values of Z_t^* gives α^{th} percentile as $Z_B^*(\alpha)$. From each repetition, $Z_B^*(\alpha)$ was then calculated for five different significance levels (α). The proposed bootstrap method of determining $Z_B^*(\alpha)$ is data based and alternatively we determine parametric simulation method, which determines $C_{\rho(0),\alpha}$. For this purpose, I have obtained the values $C_{\rho(0),\alpha}$ using each time 10000 simulated observations. The values of $Z_B^*(\alpha)$ and $C_{\rho(0),\alpha}$ are listed in Table 1.

Conclusion

- The values of $Z_B^*(\alpha)$ are reasonably close to the $C_{\rho(0),\alpha}$ for bivariate normal autocorrelated processes.
- Further, accuracy of estimates depends upon size of in control data. Clearly, the bias in estimating control limits when in control observations are 200 is very less as compared to bias when the sample size is 50.
- The bias of estimating UCL at 90% and 95% are small even for smaller in control sample sizes.
- The proposed bootstrap method of determining $Z_B^*(\alpha)$ is data based.

References

1. Alt, F.B. 1985. Multivariate quality control, Encyclopedia of statistical science, Volume 6, edited by S. Kotz and N.L. Johnson, John Wiley and Sons, New York. pp.110-122.
2. Davison, A.C. and Hinkley, D.V. 1999. Bootstrap methods and their applications, Cambridge University Press.
3. Efron, B. and Tibshirani, R. 1998. An introduction to the Bootstrap, Chapman and Hall , CRC Bocu raton London, New York, Washington, D.C.
4. Hayter, A.J. and Tsui, K.L. 1994. Identification and quantification in multivariate quality control problems. *J. Qual. Technol*. 26: 197-208.
5. Jackson, J.E. 1985. Multivariate quality control. Communications in Statistics. *Theory Meth.* 14: 2657-2688.
6. Kalgonda, A.A. and Kulkarni, S.R. 2004. Multivariate control charts for autocorrelated process. *J. Appl. Stat*. 31(3): 317-328.
7. Kramer, H.G. and Schmid, W. 1997. EWMA charts for multivariate time series. *Sequential Anal*. 16(2): 131-154.
8. Seppala, T., Moskowitz, H., Plante, R. and Tang, J. 1995. Statistical process control via the subgroup bootstrap. *J. Qual. Technol*. 27(2): 139-153.
9. Stoumbus, Z.G., Marion, R., Reynolds, Jr., Ryan, T.P. and Woodall, W.H. 2000. The State of statistical process control as we proceed into the 21st century. *J. Amer. Stat. Assoc.* 95(451): 992-998.
10. Theodossiou, P.T. 1993. Predicting shifts in the mean of a multivariate time series process: An application in predicting business failures. *J. Amer. Stat. Assoc.* 88(422): 441-447.
11. Tracy, N.D., Mason, R.L. and Young, J.C. 1992. Multivariate control charts for individual observations. *J. Qual. Technol.* 24: 88-95.