UNIVERSITEIT VAN TILBURG

Financial Time Series: Discrete or Continuous?

by
Lina Jin 176803

A thesis submitted in partial fullfilment of the requirement for the degree of Bachelor in Econometrics and Operations Research

Faculty of Economics and Business Adminstration
Tilburg University

Supervised by F.C.Dorst

Date: 27-06-2008

Contents

1 Introduction

No matter whether we have consciousness or not, time uses its invisible hand to guide our life and shape our ideas. In our daily life, we realise although not willing to accept that we are getting older, the computers we use become faster and faster, and our ipod becomes smaller and smaller and etc, etc.All these go with time. Theory finds its root in reality. Therefore, poeple introduces the time into our econometric models and we call it time series.

Time series have wide applications in economics and finance.Well-known concepts like GDP, interst rates, stock price etc can all be treated by either simple or complicated time series models. Although economic data usually have a discrete nature, econometric models generally treat them as continuous observations. Recent interest has been shown in modelling these time series with discrete models too. Generally speaking, since these economic series are positive and evolve stable around their long term mean, they can be regarded as positive, mean reverting processes no matter which model has been applied to their evolutions.In addition,since there are many similarities as well as differences among different types of models, it is the purpose of this thesis to give a comprehensive comparisons in order to figure out which model will be a suitable one for a certain type of financial or economic dataset.

The simplest mean reverting model under investigation is autoregressive process of order one with Gaussian innovations. Next to this continuous model, another autoregressive process with only positive innovation will be considered. Besides these two continuous models, a discrete time series with integer counts will be introduced too in section 2. Their basic properties will be outlined in this section and a comparison will be provided to deepen the understanding of their similarities and differences.

After the models and their properties have been described, simulation will be carried out in section 3. Two types of estimators from both ordinary least square and maximum likelihood methods will be derived and a comparison will be made. I simulate three processes by intentionally choosing the same mean and variance to see whether they will share similar evolving paths or not.

Although in simulation they didn't fail to meet with my expectation and behave quite the similar, the result from handling the real data disperses, which gives a proof that some economic data can only be analysed by certain types of models owning to the compatibility of the attributes of data and that of the models. This analysis will be carried out in section 4.

From building up the models to applying the models to real data sets, it seems that I do cover the whole research process. Unfortunately, instead of coming to an end, I am just at the very beginning. Questions still need to be answered like how to deal with the heteroskedasticity in that integer model, whether these is possiblity of introducing other types of innovations so that three models will have more equal basis in a sense of having same number of parameters, whether we can generalize the model and put them into families etc. They will be treated in a way no more than an introduction for interested readers and be handled in section 5.

2 Model Description

To begin with, I would like to present three univariate models together with their properties. In addition, a comprehensive comparison will be helpful to distwiguish one from another.

2.1 Autoregressive Model with Gaussian Innovations

It is well known that Dow Jones Industrial Average drops from 13,000 to 12,200 not in one day, but it's less well known that we may interpret today's level as that of yesterday plus a shock which might be for example an announcement of Federal Reserve about its plan to raise the interest rate level. Likewise, an autoregressive model described below can be used to model such processes. A detailed description of autoregressive models can be found in any Econometric Method textbooks like *a Guide to Modern Econometrics* from Verbeek(2004).

A time series $Y_0, Y_1 \ldots\ldots Y_n$ is a series of variables depending on its own past, which can be generated by the following model:

$$Y_t = \mu + \alpha(Y_{t-1} - \mu) + \epsilon_t \tag{1}$$

where the disturbance ϵ_t is labelled as innovation in the model, which has been generally acknowledged as a white noise process. Each element has $E(\epsilon_t) = 0, E(\epsilon_t^2) = \sigma_\epsilon^2$, and $Cov(\epsilon_t, \epsilon_s) = 0$ for all $s \neq t$. Occasionally, they may be regarded as identically independent normally distributed.

In addition, the model is called autoregressive because if Y_{t-1} and ϵ_t are uncorrelated, then

$$\mathrm{E}[Y_t\,|Y_{t-1}]{=}\mu{+}\alpha(Y_{t-1} - \mu)$$
$$\mathrm{Var}[Y_t\,|Y_{t-1}]{=}\ \sigma_\epsilon^2$$

The original autoregression of order one can be of course developed to oder p or other complicated models like moving average (MA) or autoregressive moving average (ARMA) models. But here only AR(1) will be paid attention on because of its simplicity together with its usefulness in modelling financial series.

One of the most important properties is stationarity. A stochastic process is said to be stationary given that the variances and autocovariance are finite and independent of time. If a process is stationary, then we may expect that the process varies around its mean and won't go explorsive. In this AR(1) case, the process is stationary, if $|\alpha| < 1$,ϵ_t is white noise, then

$$\mathrm{E}[Y_t]{=}\ \mu \quad \text{for all t}$$
$$\mathrm{Var}[Y_t]{=}\ \frac{\sigma_\epsilon^2}{1-\alpha^2}$$
$$\mathrm{Cov}[Y_t,\ Y_s]{=}\ \frac{\alpha^{|t-s|}\sigma_\epsilon^2}{1-\alpha^2}$$

if $|\alpha| \geq 1$, the mean,variance and covariance are not defined.

Other noteable property is autocorrelation. When stationarity is satisfied, the autocorrelation function or ACF here is defined as:

$$\mathrm{Corr}(Y_t, Y_s) = \alpha^{|t-s|}$$

The ACF is a very useful tool for describing a time series process, because it helps to determine the type of series by showing how fast it decays, whether it abruptly drops to zero at some finite lags or gradually tapers off to zero. In this AR(1) case, the ACF is to be seen declining to zero with a speed decided by α.

In order to put this AR(1) model into use, we need good estimators for α, μ and σ. This AR(1) model has an error term which is normally distributed with mean of zero and variance of σ_ϵ^2. Regardless of the exact distribution of Y_t, the OLS estimator α will be consistent and asymptoticly normal provided with stationarity and ergodicity, which leads the OLS estimation to be as good as other estimation methods. (see Greene 'Econometric Analysis' for details).

Meanwhile, if Y_t is assumed to be normally distributed, maximum likelihood estimator can be derived. The likelihood function with $\mathrm{E}(Y_t) = \mu$, $\mathrm{Var}(Y_t) = \frac{\sigma_\epsilon^2}{1-\alpha^2}$ will be:

$$L(\mu, \sigma, \alpha \,\Big|\, Y_0, Y_1 ...) = \left(\frac{1}{\sqrt{2\pi}\sigma_\epsilon}\right)^n \prod_{t=1}^{n} exp\left(-\frac{(Y_t - \mu - \alpha(Y_{t-1} - \mu))2}{2\sigma_\epsilon^2}\right) \quad (2)$$

The derived estimators which need to be sloved simultaneously will be:

$$\hat{\alpha} = \frac{\sum_{t=1}^{n}(Y_t - \hat{\mu})(Y_{t-1} - \hat{\mu})}{\sum_{t=1}^{n}(Y_{t-1} - \hat{\mu})^2}$$

$$\hat{\mu} = \frac{\sum_{t=1}^{n}(Y_t - \hat{\alpha}Y_{t-1})}{n(1-\hat{\alpha})}$$

$$\hat{\sigma}_\epsilon^2 = \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \hat{\mu} - \hat{\alpha}(Y_{t-1} - \hat{\mu})\right)^2$$

2.2 Poisson Autoregression Model

The second model is a count data time series which might for example be applied to modelling the evolution of number of injuries. The total number of injuries at time t can be decomposed into the number of remaining injuries at time t who are injuried at t-1 together with the number of new injuries arriving the pool at time t. This process can be described by a series $Y_0, Y_1 \ldots\ldots Y_n$ generated by the following model:

$$Y_t = \alpha o Y_{t-1} + \epsilon_t \quad (3)$$

where Y_0 is Poisson distributed with mean of λ_0 and ϵ_t's are independently Poisson distributed, i.e. $\epsilon_t \sim Poi(\lambda)$. The thinning operator 'o' is defined as:

$$\alpha o Y_{t-1} = \sum_{i=1}^{Y_{t-1}} B_{it}$$

where B_{1t}, B_{2t},...,for t=1,2,... are i.i.d., with

$$P(B_{it} = 1) = 1 - P(B_{it} = 0) = \alpha$$

Furthermore, B_{it} and ϵ_t are assumed to be independent. Consequently, Y_0, Y_1,... are also independent from ϵ_t and this $\alpha o Y_{t-1}$ given Y_{t-1} is $Bin(Y_{t-1}, \alpha)$ distributed. Here Y_t can be devided into two components: $\alpha o Y_{t-1}\,|Y_{t-1}$ and ϵ_t. Since the distribution of both of them is already known, the convolution of the conditional probability of Y_t given Y_{t-1} can be derived:

$$\mathbf{P}[Y_t = y_t\,|Y_{t-1} = y_{t-1}] = \sum_{s=0}^{min(y_{t-1},y_t)} \binom{y_{t-1}}{s} \alpha^s (1-\alpha)^{y_{t-1}-s} \frac{e^{-\lambda}\lambda^{y_t - s}}{(y_t - s)!} \quad (4)$$

The condtional mean and variance can be obtained:

$$E(Y_t\,|Y_{t-1}) = \alpha Y_{t-1} + \lambda$$
$$Var(Y_t\,|Y_{t-1}) = \alpha(1-\alpha)Y_{t-1} + \lambda$$

When stationarity is satisfied, Y_t will be Poisson distributed (see McKenzie(1988) for detailed proof) with

$$E(Y_t) = \frac{\lambda}{(1-\alpha)}$$
$$Var(Y_t) = \frac{\lambda}{(1-\alpha)}$$

The autocorrelation function (ACF) then will be :

$$Corr(Y_t, Y_s) = \alpha^{|t-s|} \quad \text{for k=1,2...}$$

Similarly, we may further investigate on the OLS estimators and ML estimators. Due to this thinning operator, the original model need to be reformulated as:

$$Y_t = \alpha Y_{t-1} + \epsilon_t^{'} \quad (5)$$

with a new error term to be $\epsilon_t^{'} = \epsilon_t - \lambda + \alpha o Y_{t-1} - \alpha Y_{t-1}$. The conditional mean and variance of this new error are:

$$E(\epsilon_t^{'}\,|Y_{t-1}) = 0$$
$$Var(\epsilon_t^{'}\,|Y_{t-1}) = \lambda + \alpha(1-\alpha)Y_{t-1}$$

After reformulating the original autoregressive model, the OLS estimator α is easy to be derived. Unfortunately, this OLS estimator remains consistent but not efficient any more.

Meanwhile, because the distribution of Y_t is already known, the ML estimator can be obtained as well.The first derivative of the log-likelihood function will be (see Freeland and McCabe(2004)):

$$l_\alpha = \sum_{t=1}^{n} \frac{Y_{t-1}}{1-\alpha} \frac{p(Y_t-1|Y_{t-1}-1)-p(Y_t|Y_{t-1})}{p(Y_t|Y_{t-1})}$$
$$l_\lambda = \sum_{t=1}^{n} \frac{p(Y_t-1|Y_{t-1})-p(Y_t|Y_{t-1})}{p(Y_t|Y_{t-1})}$$

Let the first derivative to be zero,the ML estimators which are not explicitly expressed can be approximated.

2.3 Autoregressive Model with Exponential Innovations

An example of the autoregressive model with exponential innovations can be found in a stochastic volatility model constructed by Nielsen and Shephard(2001). They have simulated variance of an Orstein-Uhlenbeck type driven by a *Lévy* process and this OU process can be at the same time regarded as an AR model with exponential innovations in discrete form. Interested reader may refer to this article to get an impression of the possible progress of this type of autoregressive model.

This model is again an autoregressive model of order one. A non-negative time series $Y_0, Y_1 \ldots\ldots Y_n$ is generated by:

$$Y_t = \alpha Y_{t-1} + \epsilon_t \tag{6}$$

where $Y_0 \geq 0$ has some fixed initial value and ϵ_t's are independently exponentially distributed with a positive parameter of λ, i.e.$\epsilon_t \sim \exp(\lambda)$. α here is a positive parameter too. Given Y_{t-1}, the conditional mean and variance will be:

$$E(Y_t \left| Y_{t-1}\right.) = \alpha Y_{t-1} + \frac{1}{\lambda}$$
$$Var(Y_t \left| Y_{t-1}\right.) = \frac{1}{\lambda^2}$$

Suppose that the process is stationary or $\alpha < 1$, then

$$E(Y_t) = \frac{1}{\lambda(1-\alpha)}$$
$$Var(Y_t) = \frac{1}{\lambda^2(1-\alpha^2)}$$

Meanwhile, the autocorrelation function given stationarity is:

$$Corr(Y_t, Y_s) = \alpha^{|t-s|} \quad \text{for k=1,2...}$$

Once again,I would like to find out OLS estimators and ML estimators. OLS estimator consistent but not efficient can again easliy be found through the normal regression procedure. Because the error should be non-negative, the ML estimator can be derived from the following maximum likelihood equation (see Nielsen and Shephard(1999))

$$\lambda^{-n} exp\left\{-\frac{1}{\lambda}\left(\sum_{t=1}^{n} Y_t - \alpha\sum_{t=1}^{n} Y_{t-1}\right)\right\} 1\left(min_{1\leq t\leq n}\left(\frac{Y_t}{Y_{t-1}}\right) \geq \alpha\right) \tag{7}$$

The two ML estimators can be summarized as:

$$\hat{\alpha} = min_{1\leq t\leq n}\left(\frac{Y_t}{Y_{t-1}}\right)$$
$$\hat{\lambda} = \left(\sum_{t=1}^{n} Y_t - \hat{\alpha}\sum_{t=1}^{n} Y_{t-1}\right)/n$$

Up to now, we have got an impression of three models in terms of their formulation, explaination of the error term, ACF etc. Since they look alike with each other, it will be beneficial to put these properties together. For simplicity, hereafter the model with Gaussian innovation, the one with Poisson innovation and the one with exponential innovation will be respectively named AR(1)-N, INAR(1)-P and AR(1)-E.

2.4 Similarities and Differences

These three autoregressive models of order one have something in common. In the first place, they are similar because of the way in which they have been formulated. All of them can be regarded as a process Y_t depends on its past which has some makovian property and a shock ϵ_t uncorrelated with this past Y_{t-1}.

Although three series of ϵ_t have own distinctive distributions, they can be centerized such that all of them have a mean of zero. They will be expressed in AR(1)-N, INAR(1)-P and AR(1)-E model respectively as:

$$\epsilon_t^{''} = \epsilon_t$$
$$\epsilon_t^{''} = \epsilon_t - \lambda + \alpha o Y_{t-1} - \alpha Y_{t-1}$$
$$\epsilon_t^{''} = \epsilon_t - 1/\lambda$$

Given stationarity, equation(1),(3)and(6) can be all rewritten as:

$$y_t = \alpha y_{t-1} + \epsilon_t^{''} \tag{8}$$

where $y_t = Y_t - \mu$ in AR(1)-N model, $y_t = Y_t - \frac{\lambda}{1-\alpha}$, INAR(1)-P and $y_t = Y_t - \frac{1}{\lambda(1-\alpha)}$ in AR(1)-E.

Besides they are similar in form, they have quite a lot of similarities which can be clarified with the help of Table 1.

Model	AR(1)-N	INAR(1)-P	AR(1)-E
$E(Y_t \mid Y_{t-1})$	$\mu+\alpha(Y_{t-1}-\mu)$	$\alpha Y_{t-1}+\lambda_p$	$\alpha Y_{t-1}+\frac{1}{\lambda_e}$
$Var(Y_t \mid Y_{t-1})$	σ_ϵ^2	$\alpha(1-\alpha)Y_{t-1}+\lambda_p$	$\frac{1}{\lambda_e^2}$
$E(Y_t)$	μ	$\frac{\lambda_p}{1-\alpha}$	$\frac{1}{\lambda e(1-\alpha)}$
$VarY_t)$	$\frac{\sigma_\epsilon^2}{1-\alpha^2}$	$\frac{\lambda_p}{1-\alpha}$	$\frac{1}{\lambda_e^2(1-\alpha^2)}$
$Corr(Y_t,Y_s)$	$\alpha^{\lvert s-t\rvert}$	$\alpha^{\lvert s-t\rvert}$	$\alpha^{\lvert s-t\rvert}$

Table 1: Overview of Properties

The above table tells that all of their conditional mean can be expressed as a component linear in past observation plus some constant. The conditional variances are constant in two out of three models whereas INAR(1)-P model exhibits an ARCH behaviour, which is linear in past observation with a coefficient related to the thinning parameter α. Besides these, all ACF are the same.

Meanwhile, as already described above, all three models will have an easily approached OLS estimator α, which will be consistent in all these cases. If we apply the model only to financial time series and impose stationarity, α will always lie in (0,1).

Although at the first glance three models are similar in mathematical expressions, unfortunately,they still differ from each other with regard to quite a few aspects such as the distribution of their innovations, the characteristic of the investigated time series, the estimators from maximum likelihood method and the field in which the models can be applied etc. Table 2. presents some of the differences.

Model	AR(1)-N	INAR(1)-P	AR(1)-E
Distribution of ϵ_t	$N(0,\sigma_\epsilon^2)$	$Poi(\lambda_p)$	$Exp(\lambda_e)$
Type of variable (ϵ_t,Y_t)	continuous	discrete	continuous
Range of variable (ϵ_t,Y_t)	$(-\infty,+\infty)$	$[0,+\infty)$	$[0,+\infty)$

Table 2: Overview of some differences

Finally,we may take a look at the difference arising from the asymptotic property of the ML estimators. In AR(1)-N and INAR(1)-P, estimators satisfying some regular conditions can be proven to be asymptotic normally distributed (see Freeland and McCabe(2004) and Appendix).That is :

$$\sqrt{n}(\hat{\alpha}-\alpha)\rightarrow N(0,\sigma_\alpha^2)$$

While in the last model, $\hat{\alpha}$ has another asymptotic distribution, i.e

$$P\left(\frac{n(\hat{\alpha}-\alpha)}{1-\alpha} > y\right) \rightarrow exp(-y) \text{ as } n \rightarrow \infty$$

After multipling $\sqrt{n}$, $(\hat{\alpha} - \alpha)$ in INAR(1)-P converges to some distribution instead of being degenerating. Therefore the estimator is called $\sqrt{n}$ consistent. However, the estimator in AR(1)-E is n consistent, because here $n(\hat{\alpha} - \alpha)$ converges to a certain distribution.

After getting some impression on the properties, we'd like to continue with empirical results to further exploring these models.

3 Some Examples from Literature and Simulation

Since AR(1)-N is the simplest and most widely used model in financial time series, it is available in all kinds of financial or business articles. The empirical result of this type will be omitted here.

3.1 Application of INAR(1)-P Model

One example is presented by Kurt Brannas and Shahiduzzaman Quoreshi in 2004. They use an INARMA model which can be rewritten by this INAR model to model the number of transaction of stocks per minute. In their work, estimation is carried out by seeking the conditional and feasible generalized least squares (CLS(FGLS) estimators for the standard INMA(q) model together with a GMM procedure. Standard AIC and SBIC have been used to find the lag length. Estimators from both methods have been compared in terms of fit, impact of explanatory variables on λ and on standard errors etc. Finally, conclusion has been drawn about the advantages and disadvantages in using two estimations for forecasting number of transaction.

In addition, this model has been carried out to model the monthly guest nights in a hotel(see Kurt Brannas,Jorgen Hellstrom, Jonas Nordstrom(2002)), monthly claims (R.K.Freeland(2002) etc).

3.2 Application of AR(1)-E Model

Unfortunately, there is almost no empirical result available by using this AR(1)-E model. It has been mentioned by A.Lawrance and P.Levis(1985) about their use in estimating wind velocity. Furthermore, B.Nielsen and N.Shepard mentioned the use of this model in their continuous time linear stochastic volatility models which has only been illustrated by a Gamma process in their well known paper(*Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial times*)

3.3 a Small Experiment

In this section, I would like to present an example from simulation to further compare the characteristcs of these three models and to see how close these models could be with each other. Since the three models can be formulated similarly,same mean and variance will be chozen for three models. Other parameters can be expressed in originally chozen α_0 and λ_{p0}.

$$\mu = \frac{\lambda_{p0}}{1-\alpha_0}$$
$$\sigma_\epsilon^2 = (1+\alpha_0)\lambda_{p0}$$
$$\lambda_e = \frac{1}{\lambda_{p0}}$$

The simulation procedure will be elaborated in the appendix.

In this simulation, 2500 series of each model have been generated and each series has 1000 time periods. The initial α_0 is assigned to be 0.8 and λ_{p0} to be 1.8. By applying the equations above, all series regardless of which model they belong to have an expectation and variance of 9. With a same Y_0 and all errors in different models generated from the inverse of same CDF value, these series are indeed very similar. The figure below is abridged from one simulation.

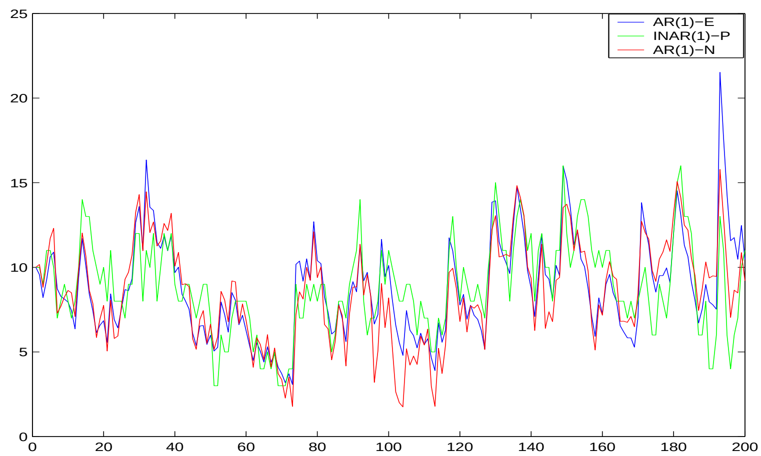

Figure 1: simulation of three series

It is not difficult to discern from the above colourful lines that they go through a lot of common ups and downs. Although sometimes one of them might deviate more than the others to give a summit, generally they exhibit a mean reverting property.

Not only the figure gives us an impression that those series are alike, but also the OLS and ML estimators of all three models will be very close to each other since they are derived from the series generated with the same means and variances. Table 3 offers an overview of all estimators from these 2500 simulations.

Method	Estimators	AR(1)-N	INAR(1)-P	AR(1)-E
OLS	$\bar{\hat{\alpha}}$	0.7957	0.7958	0.7960
	$Std(\hat{\alpha})$	0.0189	0.0191	0.0192
MLE	$\bar{\hat{\alpha}}$	0.7957	0.8006	0.8002
	$Std(\hat{\alpha})$	0.0187	0.0061	0.0002

Table 3: value of OLS and ML estimators from 2500 simulations

Just as expected, the OLS estimators $\hat{\alpha}$ in three models are very close to each other, not only in terms of their means but also the standard deviations. All these $\hat{\alpha}$ are lower than the α_0 used to generate the dataset. The ML estimators in all three models seem to perform better than their corresponding OLS estimators, not only because they are closer to the original value of 0.8, but also because they have a much smaller variance especially in INAR(1)-P and AR(1)-E models. This is of course another proof of the optimality of ML estimators.

Besides the average of OLS and ML estimators of α, I have derived value of all other estimators. For example, $\hat{\mu}$ and $\hat{\sigma}$ from OLS and ML in AR(1)-N are again very close to each other, respectively 8.999,3.2355 and 8.999,3.229. While there tends to be quite a big gap for $\hat{\lambda_p}$, with 1.8396 from OLS and 1.8035 from ML. The latter one is more accurate and closer to our original value 1.8.The same happens to $\hat{\lambda_e}$ in AR(1)-E with value of 0.5448 from OLS and 0.5741 from MLE.

4 Empirical Analysis by Applying Three AR(1) Models

After investigating on the similarity with a simulation, I am eager to know how three models will behave when dealing with the same data sets. In the coming subsections, transaction data and interest rate will be used.

4.1 Analysis of Transaction Data

The first data set that I have used is the number of transactions per minute of IBM stock in February 2005. This data set includes 19 transaction days and each day contains 390 minutes allowing for buying and selling the stocks. The summary statistics of these 19 days have been presented in the Table 4.

The averages of number of transactions per minute in all 19 days are around 41, with their maximum of 50 and minimum of 36.333. The standard deviations behave the same, with values around 17. Therefore, the number of transaction

date	Average	Standard deviation
Feb 01	39.715	18.442
Feb 02	37.741	19.846
Feb 03	40.579	16.936
Feb 04	40.492	17.753
Feb 05	34.754	14.938
Feb 06	40.115	16.829
Feb 07	41.738	18.002
Feb 08	42.990	18.189
Feb 09	39.846	14.964
Feb 10	37.982	17.234
Feb 11	46.592	18.367
Feb 12	45.926	18.837
Feb 13	40.121	18.626
Feb 14	36.605	15.236
Feb 15	50.087	18.005
Feb 16	43.636	17.777
Feb 17	41.723	18.176
Feb 18	36.333	17.07
Feb 19	41.067	16.442

Table 4: summary statistics of number of transactions

per minute from each day can roughly be regarded as independent distributed with the same distribution.

The result of average of $\hat{\alpha}$ will be presented in the Table 5:

	Estimators	AR(1)-N	INAR(1)-P	AR(1)-E
OLS	$\bar{\hat{\alpha}}$	0.3788	0.3788	0.3788
	$Std(\hat{\alpha})$	0.0845	0.0845	0.0845
MLE	$\bar{\hat{\alpha}}$	0.3768	0.3494	0.1632
	$Std(\hat{\alpha})$	0.0860	0.0407	0.0628

Table 5: value of OLS and ML estimators from transaction data

Given the assumption that Y_{t-1} and ϵ are independent in all three models, the OLS estimators will always be consist. Therefore, it is better that I derive the OLS estimators first. As expected, OLS esimators in all three models will be the same, around 0.3788. If the MLEs are consistent too, the estimated α from MLE should be round the value of OLS estimator. The results of MLE are quite controversial. First of all, AR(1)-E seems not to be a right model here, because the result of MLE is quite different from that of OLS. The test of consistency can be performed later. Secondly, We see the MLE and OLS estimator in

AR(1)-N are very close to each other in terms of both the average values of the esimators and the standard deviations. Althoug at this moement, it can not be directly concluded that this model will be the most suitable one, at least we may conclude that the AR(1)-N model can be used to analyse the process of this transaction data. Another promising result comes from the INAR(1)-P model. MLE is quite close to OLS estimator, which shows it is most likely that this MLE is consistent.In addition, the standard deviation is the lowest among all models. All in all, INAR(1)-P might be a good model to describe the process of this transaction data.

In order to test whether the OLS and ML estimators from the above three models are consistent, I have performed the Hausman Test. The detailed introduction of this test can be found in Appendix together with the adjustment I made. Because the $\widehat{\alpha}_{ols}$ is efficient in AR(1)-N model, $\widehat{\alpha}_{ols}$ and $\widehat{\alpha}_{mle}$ behave the same. Therefore, it is not necessary to apply the Hausman test for estimators derived from AR(1)-N model. Using siginficant level of 0.05,I apply Hausman Test to the other two models, resulting in the critical value of $\chi^2_{1,0.95}$= 3.8415, and $T_{INAR(1)-P}$=5.7500, $T_{AR(1)-E}$=401.0827. Quite disappointing, the consistency of average of both OLS estimator and ML estimators has to be rejected, I can only conclude that OLS estimator is consistent but MLE not. Therefore, it seems that at this moment only AR(1)-N model is the suitable model to analyse this transaction data.

4.2 Analysis of Term Structure

In this section, I will use three models to analyse data sets of term structure. The data set contains monthly interest rates taken from McCulloch and Kwon (1993).The series start in December 1946 and finish in February 1991 with total T=531. In order to adapt to the INAR(1)-P model, all data are expressed in 1:1000, which means if R1=132, then the original value is 0.132. In addition, R120 represents 10 year interest rate.

The following Table 6 gives the descriptive statistics of this data set:

Interest Rate	Average	Standard Deviation
R1	48.1940	31.9384
R2	50.0621	32.5330
R3	51.2524	32.8765
R5	52.8945	33.2358
R6	53.4727	33.4371
R11	55.0414	33.5431
R12	55.2618	33.5374
R36	58.2222	33.0427
R60	59.7326	32.6770
R120	61.5744	31.8948

Table 6: summary statistics of short term and long term interest rates

The average of interest rate increases with the increase of length of the maturity,since investors will require higher return for long term bonds than short term ones to compensate the uncertainty in a long time horizon. The volatility of bonds with different maturities remains almost the same.

I applied again three models to see whether these models will be the right ones to analyse term structure, the value of OLS and ML estimators of parameter α will be summarised below:

Interest Rate	OLS estimators	ML estimators		
		AR(1)-N	INAR(1)-P	AR(1)-E
R1	0.9803	0.9754	0.9530	0.1667
R2	0.9835	0.9785	0.9605	0.4167
R3	0.9846	0.9796	0.9625	0.5000
R5	0.9854	0.9804	0.9634	0.5714
R6	0.9853	0.9803	0.9633	0.6071
R11	0.986	0.981	0.9650	0.7143
R12	0.9864	0.9814	0.9654	0.7059
R36	0.9902	0.985	0.9731	0.7879
R60	0.9922	0.9869	0.9772	0.8387
R120	0.9943	0.9889	0.9823	0.8387

Table 7: value of OLS and ML estimators for term structure

First of all, after comparing the MLE from AR(1)-E model, I can already conclude that this AR(1)-E model won't be useful in analysing the interest rates, since there is really big gap between the OLS estimators and MLEs. The MLE of AR(1)-N won't give us much information. As the OLS estimators in AR(1)-N are efficient, I don't care really about the result of MLEs in AR(1)-N. Fortunately, it seems the values of MLEs from INAR(1)-P model are still quite close to their OLS estimators, therefore it is likely that INAR(1)-P model can be used to analyse the term structure. I am again going to use Hausman Test described in the Appendix to test the consistency of two types of estimators. Here, I don't need further adaption and just run the normal Hausman Test.

Hausman Test has been run for each serie of interest rate, the test results can be compared with the critical value of $\chi^2_{1,0.95}$=3.8415.

All of these values are larger than the critical value. Therefore, I have to reject the null hypothesis and accept the alternative hypothesis that MLEs are not consistent. This leads to the conclusion that neither INAR(1)-P nor AR(1)-E model will be suitable for analysing the term structure here.

Interest Rate	R1	R2	R3	R5	R6
TS of INAR(1)-P	11.1318	9.5725	9.4222	10.1282	10.0037
TS of AR(1)-E	9884.7	5813.4	4570.7	3585.9	2955.3
	R11	R12	R36	R60	R120
	9.6776	9.9934	9.4937	9.5865	8.8101
	1620.1	1782.7	1344	1003.2	1480.2

Table 8: result of Hausman test for interest rate

4.3 a Summary of Empirical Analysis

After analysing these two data sets, we probably realize although these three models behave very much similarly in the simulation with same mean and variance, they still have their own attributes when dealing with real data. This helps us to distinguish models, to match the models with real data and to seek for more possiblities besides the classic model with Gaussian innovations. Although I have to reject the consistency in two examples for both INAR(1)-P and AR(1)-E, I still discern that INAR(1)-P might be an option to grasp the process of number of transactions in unit period of time, since on one side I have seen the ML estimators very close to that of OLS, on the other side the Hausman Test won't fail if I just raise the significant level to 0.01, which has been also widely used in tests. In addition, AR(1)-E seems not to be an good option to analyse these discrete data sets, since its innovation varies continuously. Its MLE comes from the minimum of Y_t/Y_{t-1}, which means a single extreme value will lead the estimated value to be completely different. This is not rare to see in discrete transaction accounts. It holds the same in case of analysing interest rate which has been integerized.

Till now, I have treated two examples with the models proposed. I can't help thinking whether there are still other models as extentions to INAR(1)-P and possibly more suitable for the transaction data or even other types of financial data sets,if the INAR(1)-P is not yet completely satisfactory in analying the transaction data. The question leads to the next topic: the generalization of the AR(1) models.

5 Heteroskedasticity and Extension of AR(1) Models

5.1 Heteroskedasticity

Heteroskedasticity allowing a variation in the variance of errors is a counterpart of homoskedasticity, which violates one of the standard weak Gauss-Markov assumptions. As a result, the OLS estimators may be relatively inefficient although they may still be unbiased or consistent. In this case, generally speaking, econometricians would like to apply GLM (Generalized Linear Method) instead of OLM to capture the heteroskedasticity.

We have seen in section 2.2 that in INAR(1)-P,the conditional variance of Y_t or equivalently that of ϵ_t is not a constant but a value depending on Y_{t-1}. Obviously,this is a sign of heterskedasticity, which comes up very often in integer autoregressive models, like INAR(p) or INARMA(p) models. However,recent literature didn't show much enthusiasm about heteroskedasticity. They adopt either other approaches like GMM or just leave OLS as estimation result.

To see the difference between OLS and GLS estimators, I derive the GLS estimators by using the simulation data sets. The difference between the OLS and GLS estimators is quite trivial, the average of $\widehat{\alpha_{gls}}$=0.7960, which is sightly different from $\widehat{\alpha_{ols}}$=0.7958 in Table 3. In addition, the standard deviations lie very close too, with $Std(\widehat{\alpha_{gls}})$=0.0198 and $Std(\widehat{\alpha_{gls}})$=0.0191. Therefore, it seems the OLS is quite satisfactory in this case and there is no need to particularly derive the GLS estimators. The procedure of deriving the GLS will be found in appendix.

5.2 Extension of AR(1) Models

I have introduced in the paper three simplest AR(1) models, with classical Gaussian innovation, Poisson innovation and Exponential innovation respectively. These models can easily be extended to other members of their own distribution families. One of the benefits for such extension is that there might be less restrictions to the behavior of variables which may be reflected by the parameters. For example, by imposing the Poisson innovation, I have to assume the unconditional mean of the series equal to its unconditional variance,which is manifestly a big disadvantage of AR(1) model with Poisson distribution. This happens to AR(1)-E too, the unconditional variance is the square of unconditional mean. These restrictions probably partially explain why the last two models tend to be not useful. Another reason for such extension is that comparison might be carried out with the family such that the most suitable analyse tool can distinguish itself out which helps us to model the process more accurately.

Grunwald, Hydnman and Tedesco have summarised almost all the AR(1) models appearing in literature in (*A United View of Linear AR(1) Models*). They group all these models into classes by some characteristics of the innovations like integer variables, varying in positive real line, whole real line etc which I have also mentioned in introduction of the AR(1) models. According to their summary, a large variety of distributions come into models either through the innovations or through the series variable itself. As we already know , INAR(1)-P model is a combination of two parts, the thinning operator with Binomial distribution and the Poisson innvoation. One of the close families of Poisson innvoation is Negative Binomial with discrete nature, therefore I expect that Negative Binomial can take the place of Poisson distribution and present its own attributes. In terms of AR(1)-E model, it will be quite natural to link this Exponential distribution to Gamma distribution, since the first one is just a special case of the latter one. This extension does enable people to cope with series such as non-negative stochastic variance σ_t^2 which seems beyond the capability of AR(1)-E. As a result, we still have a lot of choices within AR(1) non-Gaussian models to deal with all sorts of financial data.

6 Conclusion

I have proposed in the first place three time series models of order one in this thesis and make a comparison of their properties from which we do see how similar these models can be. Although in simulation they still share more or less the similar sample paths, they do behave totally different when the real data sets have been used. This tells us that they are still essentially different and have their own domains for application, which helps us to build up the suitable parametric models for certain types of data.

There are still possibilities for extensions of the models. The extension of the model helps to ensure that all models will have the same number of parameters which vary more freely and all properties including the conditional expectation and variance can be set equal. With this extension, investigation can be furthered to see whether they will behave more similarly and finally share their domains of application.

7 Appendix

7.1 Deriving ML estimator in AR(1)-N

The log-likelihood function is defined as follows:
$LogL = -n(log\sqrt{2\pi} + log\sigma_\epsilon) - \sum_{t=1}^{n} \left(\frac{(Y_t-\mu-\alpha(Y_{t-1}-\mu))^2}{2\sigma_\epsilon^2} \right)$

Hereafter come the first derivative of the log-likelihood function, respectively l_μ,l_{σ_ϵ}, $\hat{\alpha}$ together with the expressions of their ML estimators.

$$l_\mu = \frac{1-\alpha}{\sigma^2\epsilon} \sum_{t=1}^{n} (Y_t - \mu - \alpha(Y_{t-1} - \mu))$$

$$\hat{\mu} = \frac{\sum_{t=1}^{n}(Y_t-\alpha Y_{t-1})}{n(1-\alpha)}$$

$$l_{\sigma_\epsilon} = -\frac{n}{\sigma_\epsilon} + \frac{1}{\sigma_\epsilon^3} \sum_{t=1}^{n}(Y_t - \alpha Y_{t-1})^2$$

$$\hat{\sigma_\epsilon^2} = \frac{1}{n} \sum_{t=1}^{n}(Y_t - \mu - \alpha(Y_{t-1} - \mu))^2$$

$$l_\alpha = \frac{1}{2\sigma_\epsilon^2} \sum_{t=1}^{n} (2(Y_t - \mu - \alpha(Y_{t-1} - \mu))(Y_{t-1} - \mu)$$

$$\hat{\alpha} = \frac{\sum_{t=1}^{n}(Y_t-\hat{\mu})(Y_{t-1}-\hat{\mu})}{\sum_{t=1}^{n}(Y_{t-1}-\hat{\mu})^2}$$

7.2 Asymptotic Attributes of Parameters in INAR(1)-P

The first derivatives regarding α and λ of log-likelihood function have been given in section 2.2. The second derivatives will be given as follows: (see Freeland and McCabe for more concise expressions)

$$l_{\alpha\alpha}=\frac{1}{\alpha^2(1-\alpha)^2} \sum_{t=1}^{n} \left\{ \frac{2\alpha^2 Y_{t-1} P(Y_t-1|Y_{t-1}-1)}{P(Y_t|Y_{t-1})} - \alpha^2 Y_{t-1} \right.$$
$$\left. + \frac{\alpha^2 Y_{t-1}(Y_{t-1}-1)P(Y_t-2|Y_{t-1}-2)}{P(Y_t|Y_{t-1})} - \left(\frac{\alpha Y_{t-1} P(Y_t-1|Y_{t-1}-1)}{P(Y_t|Y_{t-1})} \right)^2 \right.$$

$$l_{\lambda\lambda} = \frac{1}{\lambda^2} \sum_{t=1}^{n} \left\{ \frac{P(Y_t-2|Y_{t-1})}{P(Y_t|Y_{t-1}} - \left(\frac{P(Y_t-1|Y_{t-1})}{P(Y_t|Y_{t-1})} \right)^2 \right\}$$

$$l_{\alpha\lambda} = \frac{1}{\lambda\alpha(1-\alpha)} \sum_{t=1}^{n} \left\{ \frac{\lambda\alpha Y_{t-2} P(Y_t-2|Y_{t-1})}{P(Y_t|Y_{t-1})} - \frac{\lambda P(Y_t-1|Y_{t-1})\alpha P(Y_t-1|Y_{t-1}-1)}{P(Y_t|Y_{t-1})^2} \right\}$$

Given the second derivatives, the maximum likelihood estimators $\hat{\alpha}$ and $\hat{\lambda}$ jointly have the following asymptotic distribution:

$$\sqrt{n} \begin{pmatrix} \hat{\alpha}\text{-}\alpha \\ \hat{\lambda}\text{-}\lambda \end{pmatrix} \quad \Longrightarrow N\left(0, I^{-1}\right)$$

Where the matrix I is fisher's information matrix which contains the second derivatives of log-likelihood function. Furthermore, we can abridge the asymptotic distribution of $\hat{\alpha}$ from the above joint distribution, namely what we have already seen in section 2.4 :

$\sqrt{n}(\hat{\alpha} - \alpha) \qquad \Longrightarrow N(0, V)$ with V as the element at position 1×1 in I^{-1}

Or it can be rewritten as:

$$\hat{\alpha} \qquad \Longrightarrow N(\alpha, \tfrac{1}{n}V)$$

7.3 Simulation Procedure

Step One: data generation
1. Generate random variables F of $R^{1000\times2500}$ $f_i \in [0,1]$, $\forall f_i \in F$.By doing so, we get respectively the value for three errors, which satisfy $P(\epsilon_n \geq e_{in}) = 1 - f_i$, $P(\epsilon_p \geq e_{ip}) = 1 - f_i$, $P(\epsilon_e \geq e_{ie}) = 1 - f_i$.
2. Use norminv, poissinv and expinv and a given λ_{p0} to get series of three errors.
3. Give α_0, generate the Y_{0n} from nomal distribution with $\mu = \lambda_{p0}/(1-\alpha)$ and $\sigma_\epsilon^2 = \lambda_{p0}/(1-\alpha)$. Let $Y_{0e} = Y_{0n}$ and Y_{0p} be the integer part of Y_{0n}.
4. Use equation (1),(6) and given $\alpha_0$to generate series of Y_t in AR(1)-N and AR(1)-E
5. Distinguish two occasions in generating Y_t in AR(1)-P.
5.1 When $Y_{t-1} > 0$, generate $\alpha o Y_{t-1}$by using $Bin(Y_{t-1}, \alpha)$ and then use equation (3)to get Y_t
5.2 When $Y_{t-1} = 0$, let $\alpha o Y_{t-1} = 0$, then use equation (3)to get Y_t.

Step Two: OLS estimators
1. Create two series Y_t and Y_{t-1} from the origianl series Y_t. This process is conducted for all those 2500 scenarios from three types of models
2. Apply normal regression procedure for OLS estimators with Y_t dependent variable and a constant and Y_{t-1} as explanatory variables.

Step Three: Derive ML estimators
1. AR(1)-N estimators
1.1 create an error term from $\hat{\alpha} - \frac{\sum_{t=1}^{n}(Y_t-\hat{\mu})(Y_{t-1}-\hat{\mu})}{\sum_{t=1}^{n}(Y_{t-1}-\hat{\mu})^2}$ and substitute $\hat{\mu}$ with the formula $\hat{\mu}=\frac{\sum_{t=1}^{n}(Y_t-\hat{\alpha}Y_{t-1})}{n(1-\hat{\alpha})}$
1.2 Use iteration to find $\hat{\alpha}$ such that the above error term is close to zero under a certain requirement for accuracy .
1.3 Insert $\hat{\alpha}$ in the formulas for other likelihood estimators mentioned in the paper to get estimators for μ and σ.

2. AR(1)-P estimators
2.1 Create a function for conditional probability $P(Y_t\,|Y_{t-1})$.

2.1.1 Case one: when $Y_{t-1} < Y_t$, use equation 4 and let s vary from 0 to Y_{t-1}.

2.1.2 Case two: when $Y_{t-1} \geq Y_t$, use equation 4 and let s vary from 0 to Y_t.

2.2 Create a function to obtain the derivative of log-likelihood function against α

2.2.1 Compute $P(Y_t - 1\,|Y_{t-1} - 1)$ and let $P(Y_t - 1\,|Y_{t-1} - 1)$ =0, whenever $Y_{t-1} - 1 = 0$ or $Y_t - 1 = 0$

2.2.2 Compute $\frac{Y_{t-1}}{1-\alpha}\frac{p(Y_t-1|Y_{t-1}-1)-p(Y_t|Y_{t-1})}{p(Y_t|Y_{t-1})}$ and aggregate the result for 1000 observations.

2.2.3 Insert α_0 and λ_{p0} as begin value, adjust α_0 to get a l_α under required accuracy as close as possible to zero.

3. AR(1)-E estimators

3.1 Use formula $\hat{\alpha} = min_{1\leq t\leq n}\left(\frac{Y_t}{Y_{t-1}}\right)$ to get $\hat{\alpha}$.

3.2 Plug the $\hat{\alpha}$ in the formula $\hat{\lambda} = \left(\sum_{t=1}^{n} Y_t - \hat{\alpha}\sum_{t=1}^{n} Y_{t-1}\right)/n$ and $\hat{\lambda}$ can be obtained.

7.4 Introduction of Hausman Test and Its Adjusted Use

7.4.1 Hausman Test

Hausman Test is meant to evaluate the significance of an estimator against its alternatives. Now we have two different estimators $\widehat{\alpha}_{ols}$ and $\widehat{\alpha}_{mle}$ for parameter α. Suppose there exists a set S_{both} under which both estimators are consistent, with $\sqrt{n}(\widehat{\alpha}_{mle} - \alpha)$ to be efficient and $\sqrt{(n)}(\widehat{\alpha}_{ols} - \alpha)$ less efficient. Suppose at the same time there exists another set S_1 under which $\widehat{\alpha}_{ols}$ is consistent and $\widehat{\alpha}_{mle}$ not. Since these two sets are mutually exclusive, then we can construct a test: H0: both $\widehat{\alpha}_{mle}$ and $\widehat{\alpha}_{ols}$ consistent,$\widehat{\alpha}_{mle}$ efficient and $\widehat{\alpha}_{ols}$ not against H1: $\widehat{\alpha}_{ols}$ consistent and $\widehat{\alpha}_{mle}$ not. This is what Hausman Test does. With the fact that $\sqrt{(n)}((\widehat{\alpha}_{ols} - \alpha) - (\widehat{\alpha}_{mle} - \alpha))$ will be normally distributed with center of zero under H0, test statistics turns out to be:

$$\begin{gathered}(\widehat{\alpha}_{ols} - \widehat{\alpha}_{mle})'(Var(\widehat{\alpha}_{ols} - \widehat{\alpha}_{mle}))^{-1}(\widehat{\alpha}_{ols} - \widehat{\alpha}_{mle}) \\ \text{with } \sqrt{n}(\widehat{\alpha}_{ols} - \alpha) \Longrightarrow N(0, nVar(\widehat{\alpha}_{ols})) \\ \text{and } \sqrt{n}(\widehat{\alpha}_{mle} - \alpha) \Longrightarrow N(0, nVar(\widehat{\alpha}_{mle}))\end{gathered}$$

Under the null hypothesis, the test statistics will be χ^2_k distributed, where k refers to the number of estimators in the factor $\widehat{\alpha}_{ols}$, which is one here.

In addition, under H0, we can replace $(Var(\widehat{\alpha}_{ols} - \widehat{\alpha}_{mle}))^{-1}$ by $(Var(\widehat{\alpha}_{ols}) - Var(\widehat{\alpha}_{mle}))^{-1}$, so that we can easily derive the test statistics. I will not prove the above statement here, but interested readers can refer to Gelbach's lecture notes for detailed proof together with the discussion of use of this test.

As a summary, Hausman Test is formulated as:

H0: $\widehat{\alpha}_{ols}$ is consistent, $\widehat{\alpha}_{mle}$ is consistent and efficient.

H1: $\widehat{\alpha}_{ols}$ is consistent but $\widehat{\alpha}_{mle}$ not.

Test Statistics: $T = (\widehat{\alpha}_{ols} - \widehat{\alpha}_{mle})'(Var(\widehat{\alpha}_{ols}) - Var(\widehat{\alpha}_{mle}))^{-1}(\widehat{\alpha}_{ols} - \widehat{\alpha}_{mle})$, which is χ^2_1 distributed.

7.4.2 Adjusted Hausman Test

We are going to adjust the test to make it applicable in our two cases separately.First of all since we assume that the transaction data of each day will be identically independently distributed, as a result they will have the same parameter α. Therefore $\overline{\widehat{\alpha}}_{ols}$and $\overline{\widehat{\alpha}}_{mle}$ can be regarded as the estimators for parameter α too. We use $\overline{\widehat{\alpha}}_{ols}$and $\overline{\widehat{\alpha}}_{mle}$ to perform this Hausman test.

H0: $\overline{\widehat{\alpha}}_{ols}$ is consistent, $\overline{\widehat{\alpha}}_{mle}$ is consistent and efficient.

H1: $\overline{\widehat{\alpha}}_{ols}$ is consistent but $\overline{\widehat{\alpha}}_{mle}$ not.

Test Statistics: $T = (\overline{\widehat{\alpha}}_{ols} - \overline{\widehat{\alpha}}_{mle})^{'}(Var(\overline{\widehat{\alpha}}_{ols}) - Var(\overline{\widehat{\alpha}}_{mle}))^{-1}(\overline{\widehat{\alpha}}_{ols} - \overline{\widehat{\alpha}}_{mle})$, which is χ^2_1 distributed.

Second adjustment is to derive the $\overline{\widehat{\alpha}}$ for both OLS estimator and ML estimator. We have in total 19 i.i.d series of transaction data. Therefore,we can get the average and variance of the average directly:

$$\overline{\widehat{\alpha}} = \frac{1}{19}\sum_{i=1}^{19}\hat{\alpha}_i$$

$$Var(\overline{\widehat{\alpha}}) = \frac{1}{19^2}\sum_{i=1}^{19} Var(\hat{\alpha}_i)$$

Here, each $\hat{\alpha}_i$ has the same distribution described in section 7.2, namely $\hat{\alpha} \Longrightarrow N(\alpha, \frac{1}{n}V)$ with V situated at 1×1 in I^{-1}.

To perform the Hausman Test, we still need to make additional adjustment for $Var(\overline{\widehat{\alpha}}_{mle})$ in AR(1)-E. As we have seen in the introduction of this test, both of these estimators are $sqrt(n)$ consistent, therefore, we can directly use the variance of these estimators. While in AR(1)-E model, as we have seen in section 2.4, the asymptotic distribution of α is n consistent, which means this MLE converge to its mean with much faster speed than that of its OLS counterpart. As a result, the variance of MLE is so minute that we can set it to zero.

7.5 Derivation of GLS

The conditional variance in INAR(1)-P is derived as: $Var(Y_t\,|Y_{t-1}\,)$=$\lambda + \alpha(1-\alpha)Y_{t-1}$, which is equivalent to the following:

$$Var(\epsilon_t^{'}\,|Y_{t-1}\,) = \lambda + \alpha(1-\alpha)Y_{t-1}$$
$$Cov(\epsilon_t^{'}, \epsilon_s^{'}\,|Y_{t-1}, ...Y_{s-1}\,) = 0 \text{ , with } s \neq t$$

Taking $\sigma_t^2 = Var(\epsilon_t^{'}\,|Y_{t-1}\,)$ together with equation (5), we discern if we divide both sides of equation (5) by $\hat{\sigma_t}$:

$$\frac{Y_t}{\hat{\sigma_t}} = \alpha\frac{Y_{t-1}}{\hat{\sigma_t}} + \frac{\epsilon_t^{'}}{\hat{\sigma_t}}$$

equation (5) then can be rewritten as:

$$Z_t = \alpha Z_{t-1} + \nu_t \tag{9}$$

with $Z_t = \frac{Y_t}{\hat{\sigma_t}}$, $Z_{t-1} = \frac{Y_{t-1}}{\hat{\sigma_t}}$ and $\nu_t = \hat{\sigma_t}$. After this transformation, the variance of the new error term , that is $var(\nu_t) = var(\frac{\epsilon_t^{'}}{\hat{\sigma_t}}) = 1$. By doing so, we have eliminated the heteroskedasticity in the conditional variance and the estimator derived through equation (9) is GLS estimator, $\hat{\alpha}_{gls}$. In addition, the term used for transformation $\hat{\sigma_t}$ is the estimator of the conditional variance and the residual of the regression based on equation (5)can be used as an estimator for σ_t.

References

[1] R.K.Freeland and B.P.M.McCabe *Analysis of Low count time series data by Poisson autoregression*. Journal of Time Series Analysis, 2004, Vol. 25, No. 5, p. 701-722.

[2] Rong Zhu and Harry Joe *Modelling count data time series with Markov processes based on Binomial thinning*. Journal of Time Series Analysis, 2002 Vol. 27, No. 5, p. 725-738.

[3] B.Nielsen and N.Shephard *Likelihood analysis of a first-order autoregressive model with exponential innovations*. Journal of Time Series Analysis, 1999 Vol. 26, No. 6, p. 825-842.

[4] G.K.Grunwald, R.J.Hydnman and L.M.Tedesco(1996) *A unified view of linear AR(1) models*. unpublished paper.

[5] M.Verbreek *A guide to Modern Econometrics*. 2004.

[6] W.H.Greene *Econometrics analysis*. 2000

[7] J. B. Gelbach *Econ 626 Lecture Notes: the Hausman Test* 2005