Name : Sophie Poulin
Course : Time Series
Session : Fall 2015
E-mail : sophie_mariec@hotmail.com

TS Student Project: Murder Rate in Las Vegas

Introduction

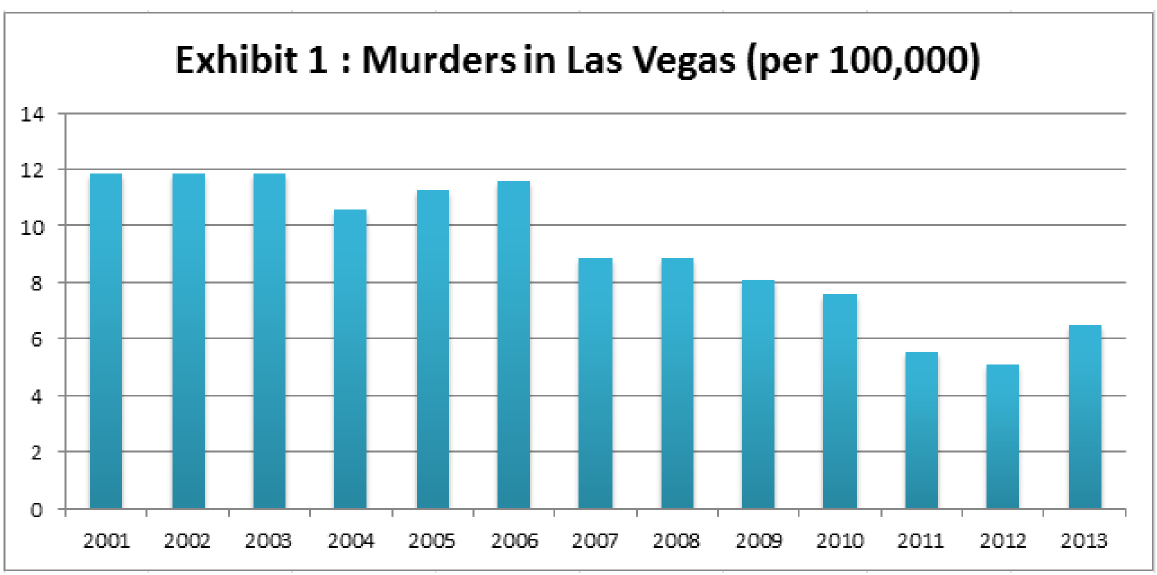
Las Vegas being one of the top tourist destinations in the world, it is renowned for its gambling, shopping, fine dining and extravagant nightlife. I thought it would be interesting to take a closer look at the crime aspect of this city. With that in mind, I decided to analyse the murder rate in Las Vegas and see if we could accurately fit the distribution to an autoregressive and/or moving average time series model. To do so, I applied statistical techniques discussed in the Time Series course.

Data

The data I used consists of murder rates in Las Vegas; number of murders per 100,000 people. The rates are annual and go from years 2001 through 2013.

Source: http://www.city-data.com/crime/crime-Las-Vegas-Nevada.html

See exhibit 1 for an overview of the data.

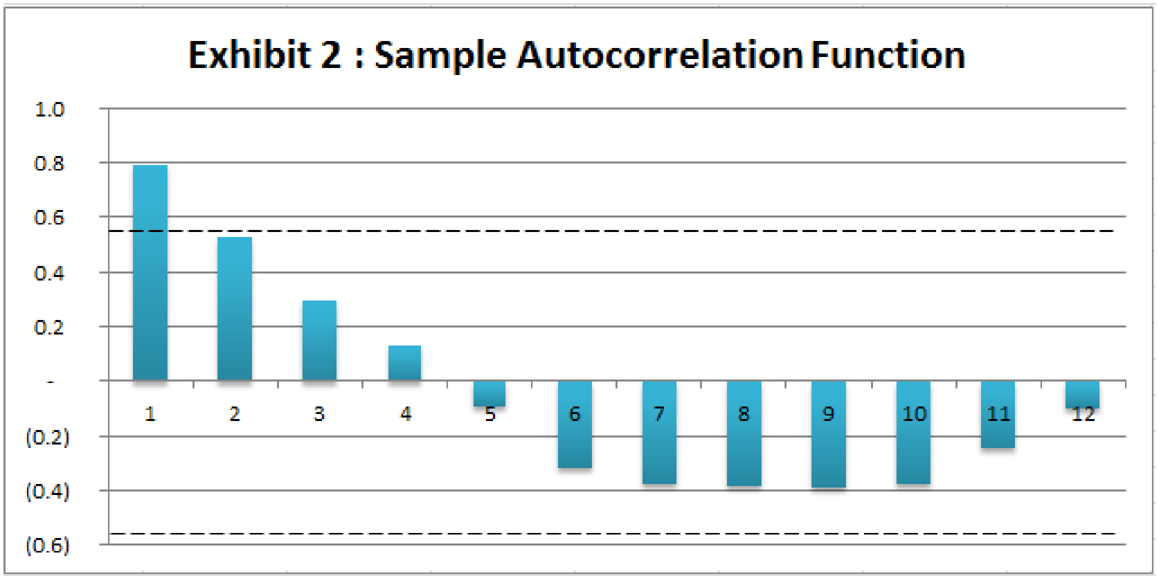


Analysis

Stationary Testing

In order to determine which model best fits our data, we will perform stationary testing. To do so, we will analyze the graph of the sample autocorrelations of the series for various lags (correlogram).

Exhibit 2 displays the sample autocorrelation function for all lags.



We observe a steady decline for smaller lags as the autocorrelation function moves toward zero.

Since all values (except for lag 1) are contained in the confidence interval of ± 0.55:

$$\pm \frac{1}{\sqrt{N}} \quad = \quad \pm \frac{1}{\sqrt{13}}$$

we can assume that our series is stationary.

Since there isn't a sudden drop in the sample autocorrelation function and it geometrically decreases up until lag 9, an autoregressive process would be more appropriate in our case as opposed to a moving average process.

First-Order Autoregressive Model – AR(1)

By using the method of moments approach, we estimate ϕ to be the lag 1 sample autocorrelation value:

$$\phi = r_1$$
$$= 0.792$$

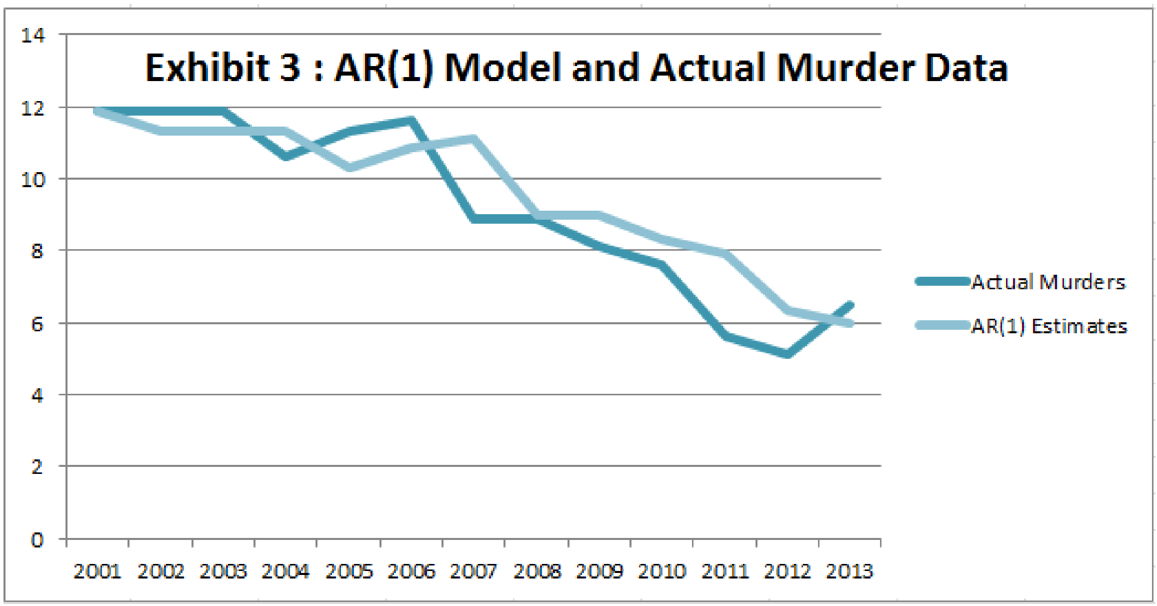
We can then easily estimate Θ_0 using the average murder rate:

$$\Theta_0 = \mu * (1 - \phi)$$
$$= 1.921$$

This leads to the following AR(1) model equation:

$$Y_t = \phi * Y_{t-1} + e_t + \Theta_0$$
$$= 0.792 * Y_{t-1} + e_t + 1.921$$

Using our estimated model we can graph the actual murder rates alongside the AR(1) estimates; see exhibit 3.



Second-Order Autoregressive Model – AR(2)

Similarly, we use the method of moments approach to estimate $\phi 1$ and ϕ_2 by applying the Yule-Walker estimates:

$$\phi_1 = \frac{r_1 * (1 - r_2)}{1 - {r_1}^2}$$

$$= 0.997$$

$$\phi_2 = \frac{r_2 - {r_1}^2}{1 - {r_1}^2}$$

$$= (0.259)$$

We can then easily estimate Θ_0 using the mean of murder rates:

$$\Theta_0 = \mu * (1 - \phi_1 - \phi_2)$$

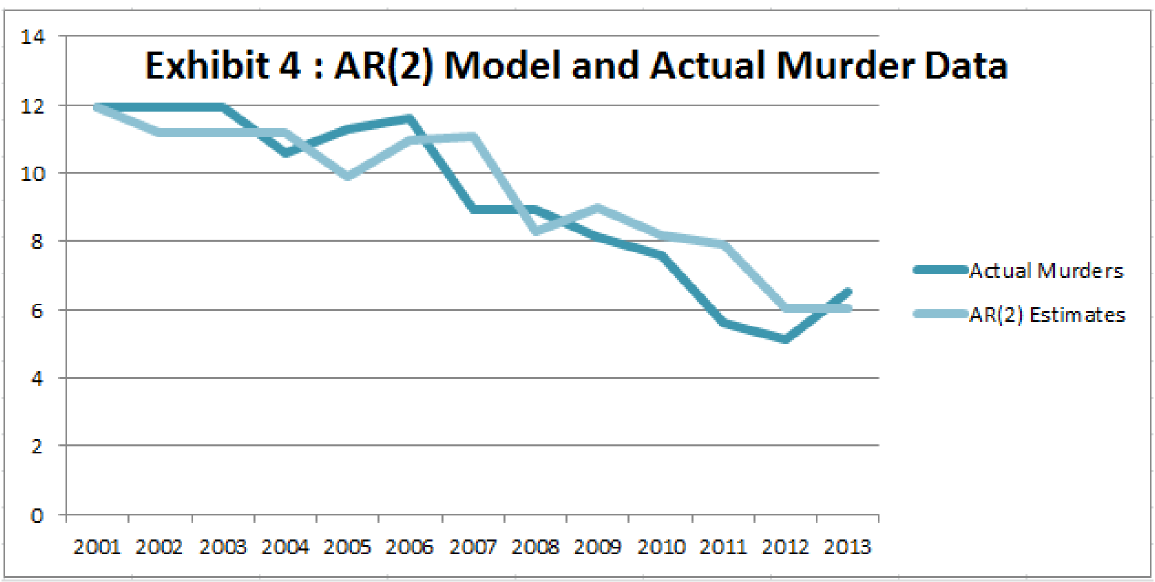
$$= 2.419$$

This leads to the following AR(2) model equation:

$$Y_t = \phi_1 * Y_{t-1} + \phi_2 * Y_{t-2} + e_t + \Theta_0$$

$$= 0.997 * Y_{t-1} - 0.259 * Y_{t-2} + e_t + 2.419$$

Using our estimated model we can graph the actual murder rates alongside the AR(2) estimates; see exhibit 4.

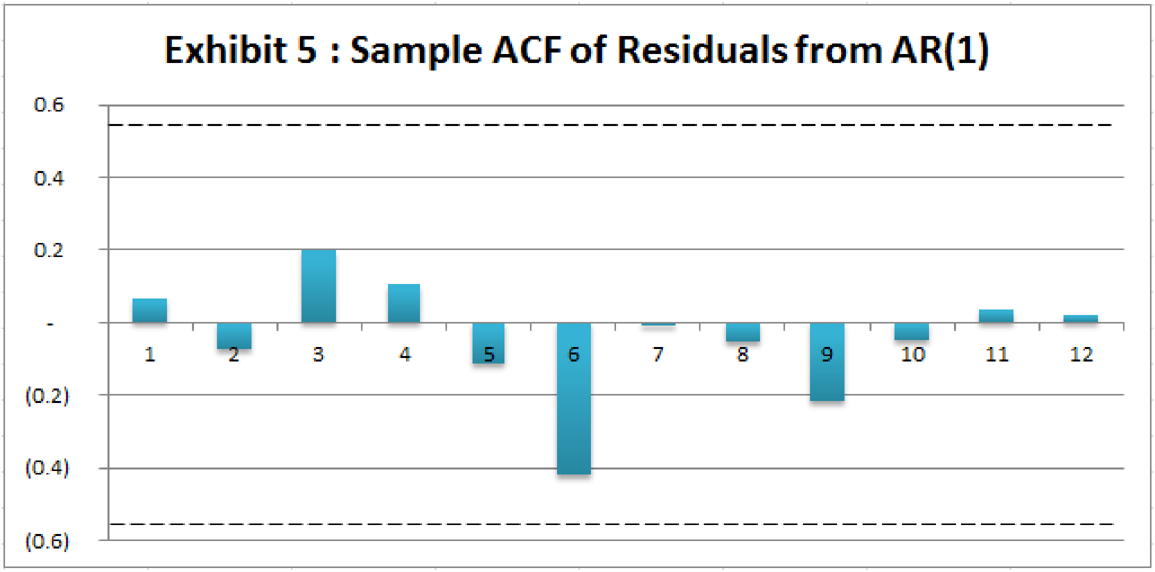


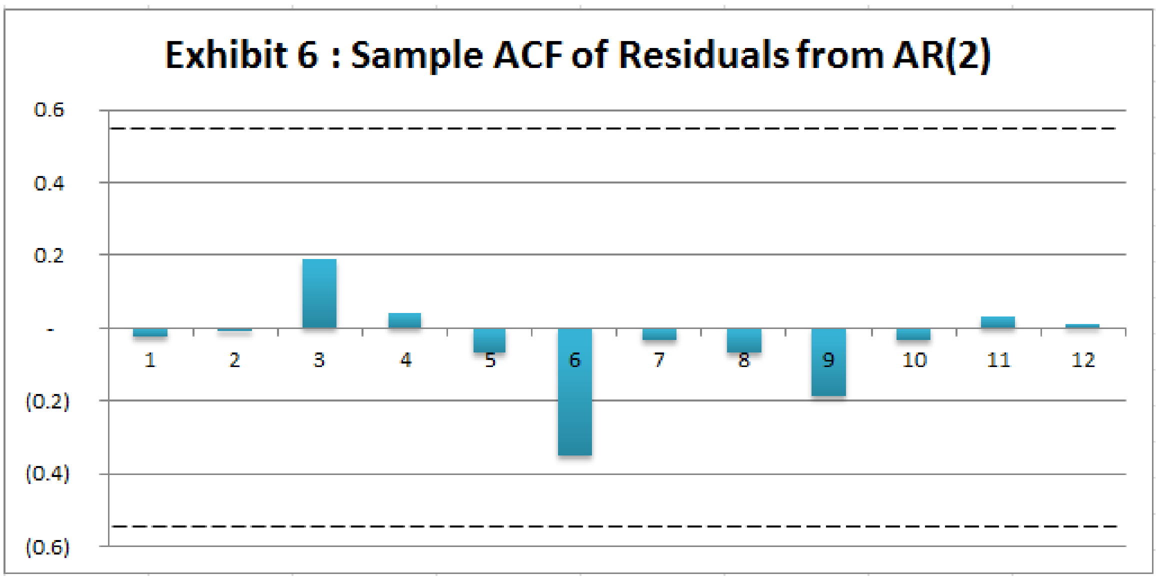
Simply by observing the AR(1) and AR(2) it is difficult to determine the most accurate model.

Residual Analysis

Since it is difficult to determine which process is more appropriate for our series, we will use residual analysis to try and get a better understanding of both models.

Exhibit 5 and 6 show the sample autocorrelation functions of residuals for both processes.





Graphs for both models lack to show statistically significant evidence of nonzero autocorrelation in the residuals since all sample ACFs are within the confidence interval.

Conclusion

Both models are plausible, and arguments could be made for either.
On the one hand the sample autocorrelation function for AR(2) is closer to 0 than for the AR(1) model.
On the other hand, if we consider the principle of parsimony, given that there is not a strong difference between the two models, we could choose the AR(1) model which is simpler.

**All detailed calculations can be found in the attached Excel Spreadsheet.*