TS Project: Annual Diameter of Skirt at hem

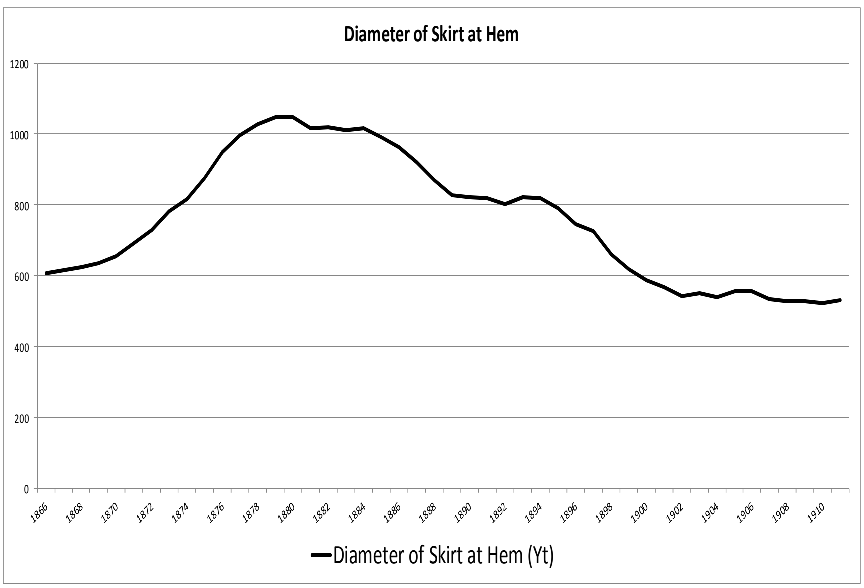
For reference, please refer to the excel file: "Time Series Project – Diameter of Skirt at Hem – Nabeel Zuberi.xlsx".

Introduction:

This project is based on the analysis of the "Annual Diameter of Skirt at Hem" data for period 1866-1911, that was pulled from datamarket.com *[Please refer to the excel file for the link to the data source]*. I will determine the model that will fit the data to predict future diameter of skirt at hem i.e. from 1911.

Data:

The data is shown in cells B14:B59 of the excel file, and represents the annual diameter of skirt at hem between 1866 and 1911 inclusive. The plot of the data can be seen in the graph below.

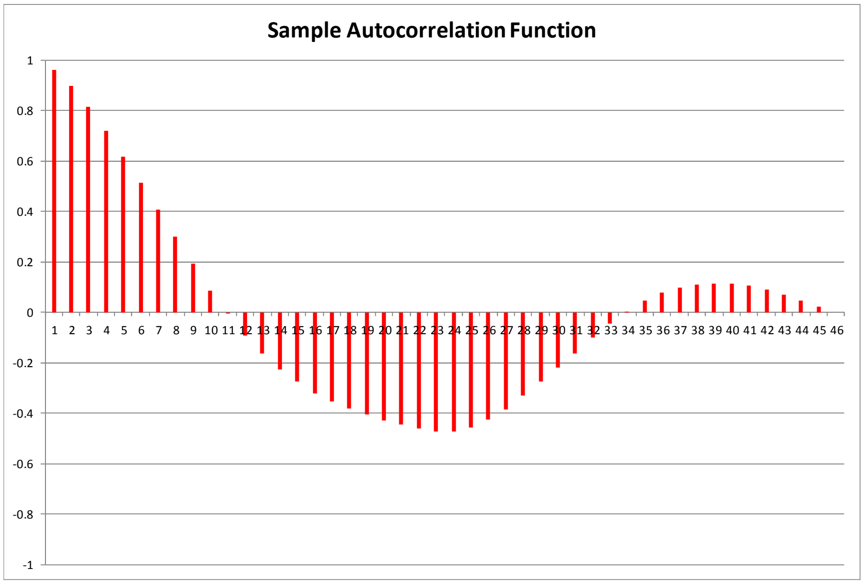


Analysis:

The plot above shows an increase in hem diameter from about 600 in 1866 to about 1050 in 1880, and that afterwards the hem diameter decreased to about 520 in 1911.

The graph of the Sample Autocorrelation is shown below *[Please refer to the excel file for the calculation]*:

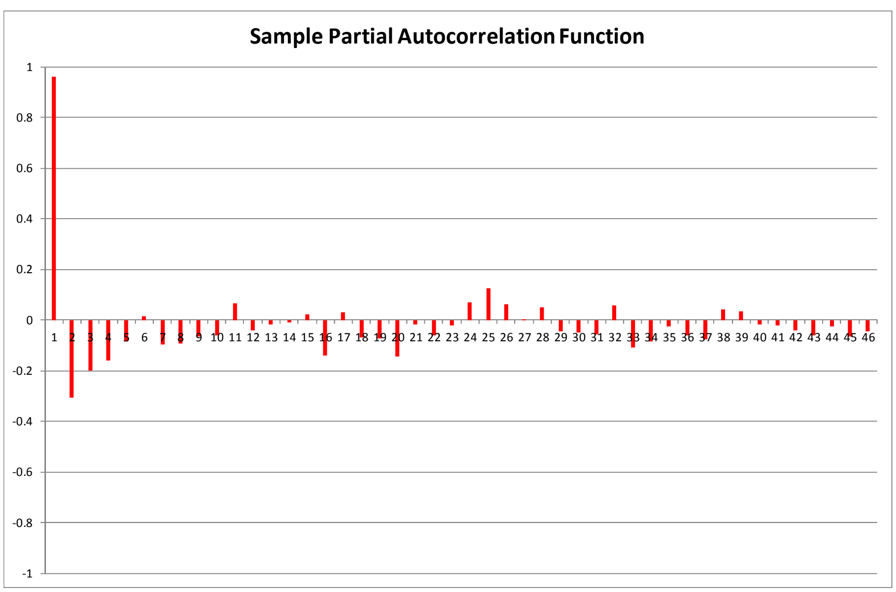
TS Project: Annual Diameter of Skirt at hem
By: Nabeel Zuberi
Session: Spring 2015



Significant autocorrelation at lag 1 of 0.9597 can be seen in the graph above which decays until lag 10 and then goes negative, increasing back to near 0 by lag 34. I will assume the data is stationary as there is little to suggest otherwise.

The lag pattern of the sample autocorrelation depicts a pattern which is more consistent with an autoregressive (AR) model as the autocorrelation decays without cutting off (as would be seen in a moving average (MA) model). Thereby, suggesting the appropriateness of an autoregressive model. I will test AR(1) and AR(2) models.

The graph of the Partial Sample Autocorrelation is shown below *[Please refer to the excel file for the calculation which was computed using formula 6.2.9]*:



The partial sample autocorrelation for lag 2 is insignificant, which suggests that an AR(2) model is likely not ideal, as an AR(p) model's PACF will cut off at lag p. Thereby, suggesting that the AR(1) model is likely to be more appropriate.

For the AR(1) model:
$\mu = 759$
$r_1 = 0.9597$

Using the method of moments gives:
$\phi = r_1 = 0.9597$.

Our AR(1) model is thus given as:
$Y_t = \mu(1- \phi) + \phi Y_{t-1} = 30.612 + 0.9597Y_{t-1} + e_t$

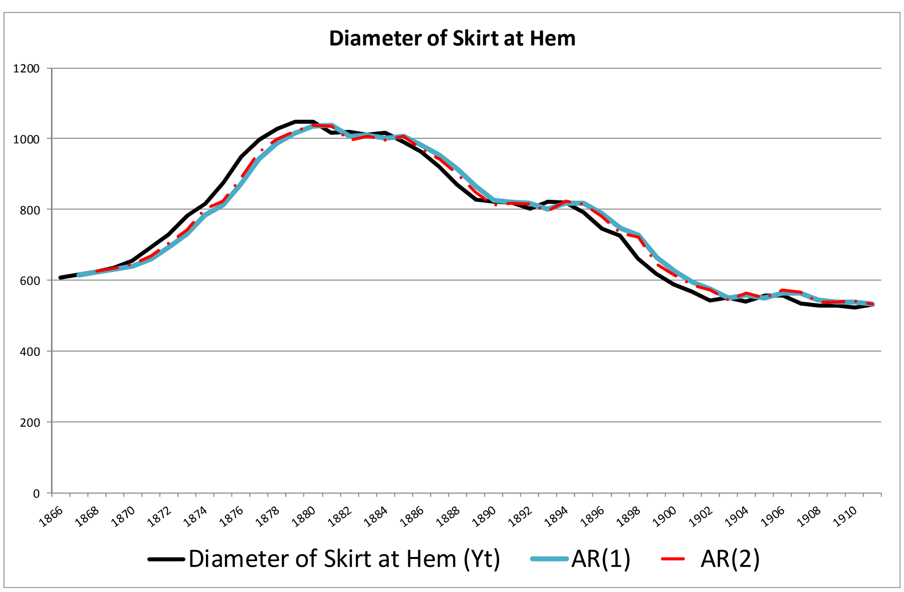
For the AR(2) model:
$\mu = 759$
$r_1 = 0.9597$
$r_2 = 0.8969$

Using the method of moments gives:
$\phi_1 = [r_1*(1-r_2)] / (1-r_1^2) = 1.2527$
$\phi_2 = (r_2- r_1^2) / (1-r_1^2) = -0.3053$.

Our AR(2) model is thus given as:
$Y_t = \mu(1 - \phi_1 - \phi_2) + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} = 39.958 + 1.2527Y_{t-1} - 0.3053Y_{t-2} + e_t$

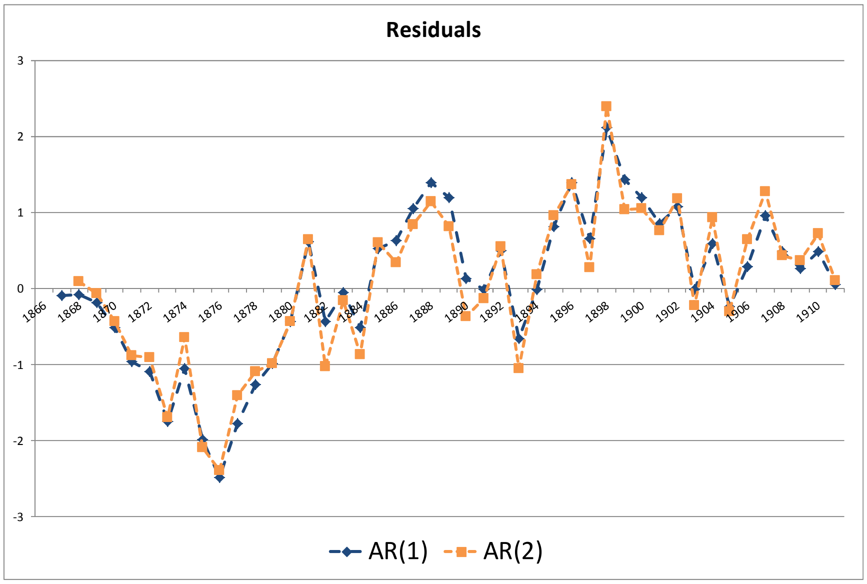
Since ϕ_2 is close to 0, which is consistent with the observation that the partial sample autocorrelation was insignificant at lag 2, this suggests that the AR(2) model does not represent a significant refinement over the AR(1) model.

The graph of the original time series with the AR(1) and AR(2) overlaid is below:



From the graph above, it is quite apparent that both the AR(1) and AR(2) models provide a reasonable fit to the historical data, but that the AR(2) model is not a significant improvement over the AR(1) model.

The graphs of the residuals for both AR(1) and AR(2) are shown below *[Please refer to the excel file for the calculation]*:



Residuals of both AR(1) and AR(2) models are similar and by the Principle of parsimony it is appropriate to select the AR(1) model since is it less complex and the AR(2) model does not present any significant improvement over AR(1) model.

Conclusion:

AR(1) model is the most appropriate choice of the model to fit this data and to predict future annual diameter of the skirts.

The final selected model is: $Y_t = 30.612 + 0.9597Y_{t-1} + e_t$