1 Theory review: MA(1)

1. Consider the process

$$X_t = W_t + \frac{1}{2} W_{t-1}$$

   (a) What is its PACF of X_t? Solve for pacf(1) and pacf(2). In the book, these are called ϕ_{11} and ϕ_{22}.

2 ACF and PACF

2. For each of the following processes,

   (i) Simulate 2000 observations with (W_t) zero-mean Gaussian white noise with variance 1.

   (ii) Plot the sample ACF. Is a clear cutoff (a lag after which all ACF values are essentially zero) present?

   (iii) Plot the sample PACF. Is a clear cutoff present?

   (a) $X_t = W_t + \frac{1}{2} W_{t-1}$;

   (b) $X_t = \frac{4}{5} X_{t-1} + W_t$;

   (c) $X_t = \frac{4}{5} X_{t-1} + W_t + \frac{1}{2} W_{t-1}$

   Hint: The function `acf2()` in the `astsa` package generates both the ACF plot and the PACF plot!

   Extra mile for the MA(1): first, add the 95% interval for MA(1) using the Bartlett's formula solved in Lab 7. Second, are the theoretical ACF and PACF values for the MA(1) calculated above close to the sample ACF and PACF values from the acf2() function?

3 Modeling

3. Think you can beat the weatherman? For this exercise, we'll be focusing on the east coast - New Haven, CT - and using ARMA methods to forecast the average yearly temperature, Y_t. To begin, download the dataset in R using: `data(nhtemp)`

   (a) Plot the data Y_t over time (this is always a good first step!).

   (b) Process the data so that it, or its residuals, is stationary. Some ideas: fit a parametric model, smooth, difference.

   (c) Plot the ACF/PACF of the resulting series from (b). Do they suggest any ARMA model?

   (d) For your time series from (b), reserve the last 10 points as a test set. Then, fit two (S)ARIMA models on the remaining data.

   (e) Generate predictions for both models using the "sarima.for" function from the "astsa" package for the test set. Which model has the lower squared error?

   (f) From the chosen model in (e), train the whole model on the entire data set and generate predictions for the 5 years.