Time series models known as ARIMA models may include *autoregressive* terms and/or *moving average* terms. In Week 1, we learned an autoregressive term in a time series model for the variable x_t is a lagged value of x_t. For instance, a lag 1 autoregressive term is x_{t-1}(multiplied by a coefficient). This lesson defines moving average terms.

A **moving average** term in a time series model is a past error (multiplied by a coefficient).

Let $w_t \overset{iid}{\sim} N(0, \sigma_w^2)$, meaning that the w_t are identically, independently distributed, each with a normal distribution having mean 0 and the same variance.

The **1st order moving average** model, denoted by MA(1) is:

$$x_t = \mu + w_t + \theta_1 w_{t-1}$$

The **2nd order moving average** model, denoted by MA(2) is:

$$x_t = \mu + w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2}$$

The **qth order moving average** model, denoted by MA(q) is:

$$x_t = \mu + w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \cdots + \theta_q w_{t-q}$$

Theoretical Properties of a Time Series with an MA(1) Model

- Mean is $E(x_t) = \mu$
- Variance is $Var(x_t) = \sigma_w^2(1 + \theta_1^2)$
- Autocorrelation function (ACF) is:

$$\rho_1 = \frac{\theta_1}{1 + \theta_1^2}, \text{ and } \rho_h = 0 \text{ for } h \geq 2$$

Note!
That the *only nonzero value in the theoretical ACF is for lag 1*. All other autocorrelations are 0. Thus a sample ACF with a significant autocorrelation only at lag 1 is an indicator of a possible MA(1) model.

For the MA(2) model, theoretical properties are the following:

- Mean is $E(x_t) = \mu$
- Variance is $Var(x_t) = \sigma_w^2(1 + \theta_1^2 + \theta_2^2)$
- Autocorrelation function (ACF) is:

$$\rho_1 = \frac{\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2}, \; \rho_2 = \frac{\theta_2}{1 + \theta_1^2 + \theta_2^2}, \text{ and } \rho_h = 0 \text{ for } h \geq 3$$

Note!
The only nonzero values in the theoretical ACF are for lags 1 and 2. Autocorrelations for higher lags are 0. So, a sample ACF with significant autocorrelations at lags 1 and 2, but non-significant autocorrelations for higher lags indicates a possible MA(2) model.

Problem 1.

Suppose that an MA(1) model is $x_t = 10 + w_t + .7w_{t-1}$, where $w_t \overset{iid}{\sim} N(0,1)$.

Find theoretical ACF and plot the same. Generate 100 observations for the above model and plot the sample data and also examine the ACF for the simulated data. Examine the above for two different sets of sample data with varying sizes.

Problem 2.

Consider the MA(2) model $x_t = 10 + w_t + .5w_{t-1} + .3w_{t-2}$, where $w_t \overset{iid}{\sim} N(0,1)$. The coefficients are $\theta_1 = 0.5$ and $\theta_2 = 0.3$.

Find the theoretical ACF for the same and plot it. Simulate 150, 200 and 250 observations for the above model and plot the data for the above model. Examine the ACF for the simulated data and compare it with the theoretical ACF