A couple comments about the problem set

$$y_t = \rho y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}$$

The mean of y_t

$$E(y_t) = \rho E(y_{t-1})$$

If $\rho \neq 1$,then the only way this can be true is if $E(y_t) = 0$.If $\rho = 1$ then E(y_t) is undefined. We assume $|\rho| < 1$.

Call $var(y_t) = E(y_t^2) = \sigma_y^2$.First, we need to compute covariances

$$\begin{aligned}
E(y_t \epsilon_t) &= E\epsilon_t (\rho y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}) = \sigma_\epsilon^2 \\
E(y_t \epsilon_{t-1}) &= E(\epsilon_{t-1} (\rho y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2})) = \rho \sigma_\epsilon^2 + \sigma_\epsilon^2 \\
E(y_t \epsilon_{t-2}) &= E(\epsilon_{t-2} (\rho y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2})) = \theta_2 \sigma_\epsilon^2
\end{aligned}$$

$$\begin{aligned}
\sigma_y^2 &= E(\rho y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2})(\rho y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}) \\
&= E[\rho y_{t-1} (\rho y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2})] \\
&+ E[\epsilon_t (\rho y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2})] \\
&+ E[\theta_1 \epsilon_{t-1} (\rho y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2})] \\
&+ E[\theta_2 \epsilon_{t-2} (\rho y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2})]
\end{aligned}$$

Forecasting rule

$$\begin{aligned}
E(y_{t+1} | y_t, \epsilon_t, \epsilon_{t-1}) &= E(\rho y_t + \epsilon_{t+1} + \theta_1 \epsilon_t + \theta_2 \epsilon_{t-1} | y_t, \epsilon_t, \epsilon_{t-1}) \\
&= \rho y_t + \theta_1 \epsilon_t + \theta_2 \epsilon_{t-1} + \underbrace{E(\epsilon_{t+1} | y_t, \epsilon_t, \epsilon_{t-1})}_{0}
\end{aligned}$$

Overlapping return horizons and predictive regression. Why we need Newey and West.

Fama: Efficient market hypothesis is that current prices reflect all publically available information. Then the return cannot be predicted based on publically available information. i.e., prices (or log prices) should follow a random walk. Let me state it like this. Let y_{t+1} be the one-period holding return on a stock or portfolio. Let I_t be a vector of all publically available information at t.

$$E(y_{t+1} | I_t) = \mu$$

In words, y_{t+1} doesn't vary with anything in I_t. Let $x_t \in I_t$ a scalar candidate information variable. We test the efficient markets hypothesis

$$y_{t+1} = \alpha + \beta x_t + \epsilon_{t+1}$$

and reject if the t-ratio on $\hat{\beta}$ is bigger than 1.96 in absolute value. If true, then y_{t+1} varies systematically with x_t.

Let's look at the S&P index, and let $x_t = d_t/P_t$ be the index's dividend yield. But instead of the one-month ahead return, I want to look at the annual return (12 month holding period).

$$y_{t,t+12} = \alpha + \beta x_t + \epsilon_{t,t+12}$$

$\epsilon_{t,t+12}$ is an MA(11) where all the MA coeffs are 1. Violates the standard assumption of iid error term in regression.

What to do econometrically?

Remember this?

$$\hat{\beta} - \beta = \frac{1}{\sum_{t=1}^{T} x_t^2} \sum_{t=1}^{T} x_t \epsilon_t$$

Assume the $x's$ are exogenous. To get the standard error of $\hat{\beta}$, we estimate

$$\text{Var}\left(\hat{\beta}\right) = \left(\frac{1}{\sum_{t=1}^{T} x_t^2}\right)^2 \underbrace{\text{Var}\left(\sum_{t=1}^{T} x_t \epsilon_t\right)}_{\text{gotta get this guy}}$$

Let $T = 4$. In the usual case, ϵ_t is iid.

$$\begin{aligned}
\text{var}\left(\sum_{t=1}^{4} (x_t \epsilon_t)\right) &= \text{var}\left(x_1\epsilon_1 + x_2\epsilon_2 + x_3\epsilon_3 + x_4\epsilon_4\right) \\
&= x_1^2\sigma_\epsilon^2 + x_2^2\sigma_\epsilon^2 + x_3^2\sigma_\epsilon^2 + x_4^2\sigma_\epsilon^2 \\
&= \sigma_\epsilon^2 \sum_{t=1}^{4} x_t^2
\end{aligned}$$

What if the $\epsilon's$ have this MA structure. We can still compute the variance, but we have to account for covariances amongst the $\epsilon's$.So

$$\text{var}\left(\sum_{t=1}^{T} x_t \epsilon_t\right) = \sum_{t=1}^{T} x_t^2 E\left(\epsilon_t^2\right) + 2\sum_{t=j+1}^{T} \left(x_t x_{t-j}\right) E\left(\epsilon_t \epsilon_{t-j}\right)$$

We could try to estimate this by

$$\sum_{t=1}^{T} x_t^2 \left(\hat{\epsilon}_t^2\right) + 2\sum_{t=j+1}^{T} \left(x_t x_{t-j}\right)\left(\hat{\epsilon}_t \hat{\epsilon}_{t-j}\right)$$

But this runs into numerical problems due to the large number of covariance terms. So Newey and West proposed making the covariances separated by long time spans less important. The Newey west estimator of this guy is

$$\sum_{t=1}^{T} x_t^2 \left(\hat{\epsilon}_t^2\right) + 2\left(1 - \frac{j}{p+1}\right)\sum_{t=j+1}^{T} \left(x_t x_{t-j}\right)\left(\hat{\epsilon}_t \hat{\epsilon}_{t-j}\right)$$

where you choose p. ($p = T^{1/4}$ is a good choice).

Lesson. In time-series regression wiht financial data, ALWAYS DO NEWEY-WEST.