Annie Pi
ECON 525 PS2
Sunday, March 2, 2025

1.

a) $\Delta yt = \alpha + \beta y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta y_{t-i} + \varepsilon_i$

After collecting data, I'd need to choose the number of p or lags, estimate the regression by OLS and obtain the t stat on beta.

b) The null hypothesis is $H_0: \beta = 0$ or that the process has a unit root and is non-stationary. The alternative hypothesis is $H_1: \beta < 0$ or that yt is stationary. The test statistic is the t-stat on $\widehat{\beta}$.

c) First we'd need to generate several simulated time series under the null and estimate the regression to compute the Dickey Fuller t-stat for each simulation. Then we can use the distribution of these simulated t-stats to determine the critical values, typically at the 5% level.

2. Unbiased forecasts mean that the current price/return reflect all information such that on average, an individual's forecast error on future price/return should be 0 aka today's price is the best forecast for tomorrow's price. Orthogonality means that forecast errors are uncorrelated with information available at the time of forecast aka you can't predict tomorrow's error using today's information because it has already been priced in.
3. Perhaps the trading strategy is exposing you to higher risk and that the high average returns are compensation for that risk instead of a violation of market efficiency. Transaction costs could also have not been considered such as bid-ask spreads, commissions, taxes, or maybe short sale constraints. Another consideration is survivorship bias because presumably we are ignoring any funds using methods that have performed poorly, inflating the alleged performance.
4. In the long run, stock returns should be positive due to the equity risk premium but the short term returns can drift unpredictably or be mean-zero. With regards to variance, returns usually show volatility clustering; variance can also be time varying from GARCH effects. The distribution can display some skewness so usually our simple normality assumptions cannot be made for modeling.
5.

a) $x_t = \alpha + \sum_{i=1}^{p} \gamma_i y_{t-i} + \varepsilon_t$ The number of lags p to include can be based on information criteria like AIC or BIC to ensure residuals are white noise. You use OLS to estimate the coefficients and you are controlling for autocorrelation in xt by including the lags of both x and y. If the coefficients are jointly statistically significant, it is implied that the past values of y contain information that helps predict x, therefore it would imply Granger causality.

b) H_0: $\gamma_1 = \gamma_2 = \gamma_3 =... \ = \gamma_p = 0$ We would use an F test to test the joint significance of the coefficients. So, we would need to estimate the 2 models, one restricted and one unrestricted. The restricted model excludes yt's lagged values. Then we would compute the sum of squared residuals for both models and plug them into the f-stat formula. We'd compare this value with the critical value from the f distribution at the specified significance level. The f-stat follows an F distribution with (p, T-k) dof.

c) Granger causality focused on prediction because it tests whether past values of one variable contain information that helps predict another. It's a test of predictive content. On the other hand, traditional causality says that changes in one variable directly causes changes in another like a true cause and effect relationship. Usually for this you would also need to consider things like endogeneity and omitted variable bias.