2016 R MFE Programming Workshop Lab 3

Brett R. Dunn

November 7, 2016

1 CAPM Failures

In this lab we are going to replicate some of basic results from Fama and French's 1993 paper *Common Risk Factors in the Returns of Stocks and Bonds*. Kenneth French provides a phenomenal data library on his website. The dataset for this week contains two files:

- *FFfactors.csv* contains returns of the famous Fama-French risk factors $mkt.RF$ (the excess return on the market), HML, and SMB along with the risk free rate RF.
- *FFports.csv* contains the returns of the 25 Fama-French portfolios. I will denote the returns of these portfolios as R_{it} for $i = 1, \ldots, 25$.

Read in both of these datasets. First we will estimate the CAPM β for each of these 25 portfolios. You will likely need to clean up the dates. Also, limit the data to be from January 1963 through the end of 2013. The β is estimated from the following time series regression for each portfolio:

$$R^e_{it} = \alpha_i + \beta_i mkt_t + \epsilon_{it} \quad t = 1, \ldots, T$$

$R^e_{it} = R_{it} - RF_t$ is the excess return on portfolio i. Now calculate the average return for each portfolio over the sample period. Plot the average return versus β_i for all 25 portfolios. If the CAPM holds, then average return should linearly increase in the β_i. Does this appear to be true?

Note doing the above will require a number of steps. To get you started, here are some hints:

- You'll need to clean up the dates so that you can subset. Use `lubridate`.
- You'll need to run regressions on groups. See the `do` verb in the `dplyr` reference manual and the associated examples. Another package called `broom` and its function `tidy` makes this particularly simple.
- You may need to use `tidyr` functions.
- You'll need to use a join or a merge. Recall that `dplyr` has some nice join functions.

2 Fama-French Model

Now we will look at the famous Fama-French 3 factor model. Instead of estimating just β_i, estimate β_i, h_i, and s_i for each portfolio using the following time series regression:

$$R^e_{it} = \alpha_i + \beta_i mkt.RF_t + h_i HML_t + s_i SMB_t + \epsilon_{it} \quad t = 1, \ldots, T$$

Calculate the average returns for the 3 Fama-French factors $E[mkt.RF_t]$, $E[HML_t]$, and $E[SMB_t]$. Now for each portfolio, calculate the predicted value:

$$pred_i = \beta_i E[mkt.RF_t] + h_i E[HML_t] + s_i E[SMB_t]$$

Plot this predicted value versus the average excess return for each portfolio. Do things look a little better?