Problem Set 2

Goal: Explore OLS & Multiple Hypothesis Testing
Application: CAPM

In this problem set, we will use real financial data to perform some statistical exercises. The main variables in the data set are the monthly returns of some stocks and financial assets, as well as the returns of the S&P 500 index, which is a measure of the returns of the stock market as a whole. We denote the S&P 500 returns as r_M.

We will analyze our stock data through the lens of an important model in Financial Economics, the CAPM model. This model states that the expected rate of return of asset j in excess of the risk-free return is determined by how its returns covary with the market return, in particular,

$$E(r_j) - r_f = \beta_j(E(r_M) - r_f) \tag{1}$$

where r_j is the rate of return of asset j, r_f is the rate of return of a risk-free asset (i.e. an asset that always yields the same return, thus taken to be a constant in CAPM theory), r_M is the market return, and $\beta_j = \frac{cov(r_j, r_M)}{var(r_M)}$ is a measure of how asset j's returns covary with market returns. We refer to $r_j - r_f$ and $r_M - r_f$ as excess returns (returns in excess of the risk-free return). In CAPM, $E(r_M) > r_f$, and thus equation 1 predicts that assets with higher β_j will have higher $E(r_j)$.

Part 1. CAPM – β and Understanding OLS

What is the intuition behind the CAPM theory? A stock that tends to vary inversely with the market allows its owners to diversify risk away from the general market. Investors will thus be willing to buy such stocks, even if expected returns are relatively low. On the other hand, a stock that strongly covaries with the market offers little risk-diversification benefit. Investors will thus require a relatively high return to buy it. In equilibrium, prices will adjust such that the high-β stocks that provide little insurance against market-wide risk will have high excess returns, and low-β stocks that better insure against market-wide risks have lower expected excess returns.

In Part 1, we will use `data_part1.dta`, and look at the market returns, r_M, as well as the the returns of two stocks: SPDR Gold Shares (GLD) and Morgan Stanley (MS) which we will call r_A and r_B, respectively.

1. Use `read.dta13` to load `data_part1.dta`. Assume the risk-free rate is $r_f = 0.41\% = 0.0041$ per month. Generate excess returns variables for the market, SPDR Gold Shares (GLD) and Morgan Stanley (MS), i.e., create new variables equal to $r_M - r_f$, $r_A - r_f$ and to $r_B - r_f$.

2. Compute the sample variance of (r_A, r_B, r_M), and separately of $(r_A - r_f, r_B - r_f, r_M - r_f)$. How do the two sample variance matrices compare? Explain the connection between the two sample variance matrices, using that r_f is a constant.

3. Install and load the `ggplot2` and `gridExtra` package. Create two scatterplots using `ggplot`, one scatterplot with $r_A - r_f$ on the vertical axis and $r_M - r_f$ on the horizontal axis, the other scatterplot with $r_B - r_f$ on the vertical axis and $r_M - r_f$ in the horizontal axis. To make the two scatterplots easier to compare, force the two scatter plots to have the same limits for the y-axis using the `ylim` option for `ggplot`, choosing limits that are large enough to include all points for both plots. Save the scatterplots as `plot1` and `plot2`. Use the command `grid.arrange(plot1, plot2, ncol=2)` to plot the two scatterplots side by side. Do the excess returns of the gold shares or of Morgan Stanley seem to be more strongly associated with market returns? Can you provide an explanation?

4. Consider the following linear regression specification:

$$(r_{jt} - r_f) = \alpha_j + \beta_j(r_{Mt} - r_f) + \epsilon_{jt} \tag{2}$$

where r_{jt} and r_{Mt} are the returns on asset j and on the market in period j, and where α_j and β_j are defined by the linear projection of $(r_{jt} - r_f)$ on a constant and $(r_{Mt} - r_f)$.

   (a) Show that β_j defined by linear projection in equation 2 equals $\frac{cov(r_j, r_M)}{var(r_M)}$. (Hint: one way to do this would be to derive the projection coefficient for $Y = r_j - r_f$ and $X = r_m - r_f$. An easier way to do this would be to start with $cov(r_j - r_f, r_m)$, substitute expression (2) for $r_j - r_f$, use the fact that r_f and a_j are constants to simplify, and then use the fact that in a general linear projection model $cov(e, X) = 0$.)

   (b) Show that, if equation 1 holds, then $\alpha_j = 0$. (Hint: take expectations on both sides of Equation (1) and use the fact that in linear project $E[e] = 0$.) Note: this implies that if $\alpha_j > 0$, the asset has a higher expected return than should be possible under CAPM, and if $\alpha_j < 0$, the asset has a lower expected return than should be possible under CAPM.

   (c) Estimate equation 2 by OLS regression separately for gold and Morgan Stanley shares using the `lm` function, then answer:

      i. What are your estimated values $\hat{\beta}_A$ and $\hat{\beta}_B$?

      ii. Are your estimated values $\hat{\beta}_A$ and $\hat{\beta}_B$ consistent with what you would expect these coefficients to be based on your estimated sample variances and covariances from question 2?

      iii. Based on the estimated $\hat{\beta}_A$ and $\hat{\beta}_B$, which of the two stocks seems to covary more closely with the market?

      iv. Produce the same scatter plots as in Questions 3, but now also overlay a regression line on each scatterplot by using `geom_smooth` with option `method="lm"`, and with the option `SE=FALSE`.

      v. Based on your results, discuss whether an investor worried about market volatility should add SPDR Gold Shares (GLD) to her portfolio or add Morgan Stanley (MS) shares to her portfolio.

Part 2. CAPM –Multiple Hypothesis Testing and the Search for α

In this part we consider the parameter α in equation 2. As evidenced by equation 1, CAPM predicts that $\alpha = 0$. We will use `data_part2.dta`, and look at the returns of 10 mutual funds – all of whom made it to Time's 50 Best Mutual Funds in 2018[1]. We will fit the CAPM regression model of equation 2 separately for each mutual fund. Note that if mutual fund j has $\alpha_j > 0$, then its expected return is higher than should be possible under CAPM. In finance, "searching for α" is searching to find such assets, or, in this case, mutual fund managers who systematically do better than should be possible under CAPM. Conversely, any mutual fund with $\alpha_j < 0$, then its expected return is lower than should be possible under CAPM, and an investor would naturally wish to avoid such funds/fund managers.

1. Use `read.dta13` to load `data_part2.dta`. Redefine all return variables to be relative to the risk free rate (i.e., subtract 0.0041 from each return variable).

2. Estimate α and β using the regression specified in equation 2 for mutual fund SWPPX.
   (a) Interpret in words the meaning of homoscedasticity versus heteroscedasticity in the context of equation 2. Give an argument for why homoscedasticity might not be plausible.
   (b) Using heteroscedastic robust standard errors, test the null hypothesis that $\alpha = 0$ against the two-sided alternative $\alpha \neq 0$ at the 0.10 significance level. Do you or do you not reject the null hypothesis?
   (c) Using heteroscedastic robust standard errors, construct a 90% confidence interval around α.
   (d) What does your results imply about mutual fund SWPPX?

3. Estimate α and β using the regression in 2 separately for each of the 10 mutual funds in the data. For each of the ten funds, test the null hypothesis $H_j : \alpha_j = 0$ against the two-sided alternative $\alpha_j \neq 0$ at the 0.10 significance level. How many null hypotheses would you reject? Would you conclude that some mutual funds systematically have higher or lower return than should be possible under CAPM? Is there any fund which you are surprised made the list for Time's 50 Best Mutual Funds in 2018?

4. In question 3, you tested 10 null hypotheses, $H_j : \alpha_j = 0$ versus $\alpha_j \neq 0$ for $j = 1, 2, .., 10$. Such a problem is called a multiple hypothesis testing problem, as you are testing multiple null hypotheses. One worry is that, since you are testing multiple null hypotheses, you may have an unacceptably high probability of incorrectly rejecting one or more true null hypothesis. Consider the **Family Wise Error Rate** (FWER), which is defined as the probability of incorrectly rejecting one or more true null hypothesis.
   (a) Explain why, if only one null hypothesis is true, the FWER of you testing procedure from question (3) is 0.10.
   (b) Explain why, if two of the ten null hypotheses are true, the FWER of your testing procedure from question (3) is greater than or equal to 0.10, and give the exact

[1] http://time.com/money/5090045/best-mutual-funds-2018/

value of the FWER if the test statistics for the two true null hypotheses are independent of each other.

(c) Explain why, if k null hypotheses are true, where k is any number from 1 to 10, the FWER of your testing procedure from question (3) is greater than or equal to 0.10. Give the exact value of the FWER if all ten of the null hypotheses are true and all ten of the test statistics are independent of each other. (See hint at end of problem set).

5. A testing procedure is said to control the FWER at level a if the probability of one or more false rejection is at most a. There are many procedures that have been developed to control the Family-Wise Error Rate, starting with the Bonferroni correction developed by Olive Jean Dunn in the late 1950s. The Bonferroni correction works as follows: instead of rejecting each null hypothesis H_j for which $p_j < a$, reject each null hypothesis H_j for which $p_j < \frac{a}{m}$ where m is the number of hypotheses in the family (in Question 3, with 10 mutual funds, $m = 10$). How many of the 10 null hypotheses that you tested in Question 3 would you reject with the Bonferroni correction?

6. The Bonferroni correction reduces statistical power substantially. The Holm or Holm-Bonferroni Step-Down Method, developed by Sture Holm in 1979, is an alternative method that also controls the FWER while rejecting at least as many null hypotheses as with the Bonferroni correction. This implies that it controls the FWER while having at least as much power to correctly reject false null hypotheses as the Bonferroni correction. The Holm procedure that controls the FWER at level a works as follows.

   (a) Rank your p-values from smallest to largest. Suppose there are m hypotheses. Let k index the rank.

   (b) Start with $k = 1$.

   (c) Compute $\frac{a}{m+1-k}$. This is your critical value for rank k. If p_k is less than this critical value, reject hypothesis k. Otherwise, do not reject hypothesis k.

   (d) If hypothesis k was rejected, repeat step (c) for $k + 1$. If hypothesis k was not rejected, the process ends here, and all other hypotheses are not rejected.

   Implement Holm's procedure for the 10 hypothesis tests that you performed in Question 3, using the p-values you computed in Question 3. How many null hypotheses do you reject? Were you able to reject at least as many null hypotheses as in question 6 using the Bonferroni correction? Were you able to reject more? (*Hint:* you can, if you wish, program the above loop, or alternatively use the `R` function `p.adjust` with method="Holm").

Hint For Part 2, Question 4(c)

As a general rule, for any two events, A and B,

$$Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]. \tag{3}$$

In the special case where A and B are independent and $\Pr[A] = \Pr[B] = p$, equation 3 simplifies to

$$Pr[A \cup B] = 2p - p^2 = 1 - (1-p)^2. \tag{4}$$

For any sequence of events $A_1, ..., A_K$, we can iterate on those rules. In the special case where $\Pr[A_1] = \Pr[A_2] = \ldots = \Pr[A_K] = p$ with $A_1, ..., A_K$ mutually independent, equation 4 generalizes to

$$Pr[\cup_{k=1}^{K} A_k] = 1 - (1-p)^K. \tag{5}$$