• You can not get a check without answering all the discussion questions.

• The report must be printed and very clear, with a professional appearance, tables and figures numbered as asked, and put at the end of the report.

Problem 1: Financial Data Analysis

On Ken French's web site, Research Returns section, get the daily and monthly files for size-sorted portfolios. Get the Fama / French Factors daily and monthly returns. From these files, get the daily and monthly returns for the lowest, and highest deciles, and the market portfolio. You now have 3 returns series daily and monthly. Keep the returns from 1931 to 2012. Look at the mean of these returns and find out if they are written in percentage returns or as raw returns. Convert them into raw returns (by dividing by 100) if needed. Then convert these discrete returns in log (continuous) returns, now work with the log-returns.

1. For each portfolio return, put in Table 1, the annualized sample mean, annualized standard deviation, the skewness, the kurtosis, the Jarque-Bera (JB) test, the first order autocorrelation of returns and of **squared** returns. Put the 6 returns in rows and the statistics in columns.

   Indicate in the text the approximate standard error and the distribution of the skewness, kurtosis, JB test, and first ACF. They will be different for the monthly and daily data.

   Put a star next to the numbers in the table that are statistically significant at the 5% level.

2. For the small decile, Compute the t-statistic for the null hypothesis that the mean monthly return was 10% per year (0.10/12 per month). Can you reject the null at the 5% level? You may think that you can write a more powerful test by using the daily

data which have more observations.

Use the small decile daily returns to compute the equivalent t-statistic for a test of the null that the mean daily return was equal to 10% annualized, that is 0.10/252 daily. Compare the two t-statistics and conclude !

3. Compare skewness, kurtosis, JB tests for daily vs monthly returns. What can you say about the normality of monthly vs. daily returns for these portfolios? Recompute Skewness, Kurtosis and JB, with only the data from 1989 to 2012, report in Table 2. Are your conclusions different?

4. Given the autocorrelations in Table 1, what can you say about the predictability of daily and monthly returns? Given $\rho(1)$ in the table, what fraction of the variance of each return can you hope to "predict" in an AR(1) regression ? Hint: What is the R-square of the regression: $r_t = \alpha + \rho r_{t-1} + e_t$?

   How many observations would you need for a first order autocorrelation estimate of 0.15 to be statistically significant at the 5% level. Consider the small firm portfolio daily return. You want to check if the predictability is reliable and stable. First, estimate the first order autocorrelation with daily returns using two years of data, 1931-1932. Then move the data by one year and recompute.

   Plot in Figure 1 the time series of the autocorrelation estimates. Compare to the average found for the whole sample. Conclude.

5. Compare the annualized standard deviations of the daily and monthly returns. Are they equal? Compute and report in the text the ratios of annualized monthly variance to

annualized daily variance for the three portfolios. Are they equal to, larger or smaller than 1?

In either of these three cases, what can be the reason? Hint: Look up variance ratio tests in Campbell, Lo, McKinlay or in articles, and think how the aggregation formula may fail.

6. Plot in Figure 2, the Cross-Correlation Function (ccf in R) of the small and large portfolios monthly returns, plot only up to 5 lags. Do you notice any lead-lag relationship one way or the other ? Can one portfolio be used to predict the other?

Problem 2: Delta Method

Consider a simple multiple regression by MLE as seen in class. We wrote the MLE of σ, the standard deviation of the regression noise. Use the Delta Method to compute the variance of the MLE of σ^2.

Problem 3: Cycle and period of a cycle

The dataset **RGDPA.txt** contains the annual real GDP growth rates for the last 77 years. For this problem, you can use either the arima, lsfit, or lm commands for these questions, and get the covariance matrix of the parameters from the output.

1. What are the mean and standard deviation of the US Real GDP growth rate over the 77 years? In Figure 3 and 4 plot the ACF and the PACF of the GDP growth rate. Is GDP growth rate predictable, which model do you think would work best?

2. Using only the data until 1987 included, estimate an AR(2)

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \sigma\epsilon_t, \quad \epsilon_t \sim N(0,1)$$

by classical methods, where $y_t = \log GDP_t - \log GDP_{t-1}$. In Table 3, give the point estimates of the coefficients and their asymptotic standard errors.

Find the asymptotic distribution of the MLE of the determinant of the characteristic equation, its point estimate and standard deviation by the Delta method.
What is the probability of a cycle given this asymptotic distribution ?

3. Is there a cycle **at** $(\hat{\phi}_1, \hat{\phi}_2)$? If yes, give the asymptotic distribution of the period of the cycle, mean and standard deviation, and a 95% confidence interval.

Problem 4: Bayesian regression and forecasting, Direct Monte Carlo

1. Continue with the GDP data and your previous estimates for the AR(2).
   If you had estimated the AR(2) by Bayesian regression with diffuse priors on the parameters, would your numerical results have been different?
   Would your interpretation of the numerical results have been different?

2. Now you will simulate the **predictive** distribution of $y_{1988,1989,1990}$. Describe your Monte Carlo algorithm, explaining the steps needed to make one draw of $y_{1988,1989,1990}$. Make 10000 draws.
   In Figure 5, give a time series plot with the actual $y_{1988}, \ldots, y_{1990}$ and, superimposed, the predictive means of the forecasts and their 25^{th} and 75^{th} percentiles.
   On Figure 6, plot the histograms of the 3 predictive densities. Are they normally distributed?
   Compute and report in Table 4 the actual growth rates for these years, your predictive means and the interquartile range of the growth rates predictive densities, as well as the predictive probability of a recession (y<0), for each year.