F500 Problem Set 2 - Solutions

Dr Tom Auld*

February 24, 2019

Question 1

a) Mathematically:

$$
\begin{aligned}
VAR(P(w)) &= VAR(wX + (1-w)Y) \\
&= VAR(wX) + VAR((1-w)Y) + 2COV(wX, (1-w)Y) \\
&= w^2 VAR(X) + (1-w)^2 VAR(Y) + 2w(1-w)COV(X,Y) \\
&= w^2\sigma_X^2 + (1-w)^2\sigma_Y^2 + 2w(1-w)\sigma_X\sigma_Y\rho_{XY} \\
&\leq w^2\sigma_X^2 + (1-w)^2\sigma_Y^2 + 2w(1-w)\sigma_X\sigma_Y \qquad \because \rho_{XY} \leq 1 \\
&= [w\sigma_X + (1-w)\sigma_Y]^2 \\
&\leq \max(\sigma_X, \sigma_Y)^2 = \max(\sigma_X^2, \sigma_Y^2)
\end{aligned}
$$

$w\sigma_X + (1-w)\sigma_Y \leq \max(\sigma_X, \sigma_Y)$ because this is a weighted sum of the two standard errors whose weights are positive and sum to unity.

Intuitively: (Without leverage) it is impossible to create a portfolio of assets whose variance is greater than all of the individual components.

b)

Assumptions: We assume investors are risk averse and thus prefer lower variance.

Comment: There is no trade off between risk and reward in this very simple setup as both assets have the same mean. In terms of portfolio an efficient frontier does not exist (or is a single point). The mean cannot be changed so it makes no sense to perform the optimisation: maximise the mean for a given amount of variance. Mathematically:

$$
w_{opt} = \underset{w \in [0,1]}{\arg\min}\, VAR(P(w))
$$

$$
\begin{aligned}
VAR(P(w)) &= w^2\sigma_X^2 + (1-w)^2\sigma_Y^2 + 2w(1-w)\sigma_X\sigma_Y\rho_{XY} \\
&= w^2\sigma_X^2 + (1-2w+w^2)\sigma_Y^2 + 2(w-w^2)\sigma_X\sigma_Y\rho_{XY} \\
&= w^2(\sigma_X^2 + \sigma_Y^2 - 2\sigma_X\sigma_Y\rho_{XY}) + 2w(\sigma_X\sigma_Y\rho_{XY} - \sigma_Y^2) + \sigma_Y^2 \\
&= w^2(VAR(X-Y)) + 2w(\sigma_X\sigma_Y\rho_{XY} - \sigma_Y^2) + \sigma_Y^2 \\
&= Aw^2 + Bw + C
\end{aligned}
$$

*Comments and corrections to tja20@cam.ac.uk

with

$$A = VAR(X-Y),\ B = 2(\sigma_X\sigma_Y\rho_{XY} - \sigma_Y^2),\ C = \sigma_Y^2$$

Thus variance is a quadratic in w. In this case the A parameter is positive which means it has a minimum. This can easily be found either my differentiation or (my favourite) by completing the square and is $-\frac{B}{2A} = \frac{(\sigma_Y^2 - \sigma_X\sigma_Y\rho_{XY})}{VAR(X-Y)}$. However, I do not actually agree with the answer in the question as we have imposed $w \in [0,1]$ and the constraints (ie endpoints of $[0,1]$) could bind. Thus, to be absolutely correct:

$$w_{opt} = \begin{cases} 0 & \sigma_Y \leq \rho_{XY}\sigma_X \\ 1 & \sigma_X \leq \rho_{XY}\sigma_Y \\ \frac{(\sigma_Y^2 - \sigma_X\sigma_Y\rho_{XY})}{VAR(X-Y)} & \text{otherwise} \end{cases}$$

A simple proof by counter example is enough to demonstrate $Var(P(w_{opt})) \leq \min(\sigma_X^2, \sigma_Y^2)$. Since $VAR(P(w))$ takes the values σ_X^2 or σ_Y^2 when $w = 0$ or 1 supposing $Var(P(w_{opt})) > \min(\sigma_X^2, \sigma_Y^2)$ contradicts the fact that $w_{opt} = \arg\min_{w\in[0,1]} VAR(P(w))$. Thus the result must hold.

c)

Flippant Answer: Never, as the question imposes $w \in [0,1]$, ie no shorting.

Short Answer: Negative weights are assumed when one asset is much more volatile than the other (ie when the ratio of standard errors exceeds ρ_{XY}^{-1}) *and* there is no cost involved in shorting.

Long Answer, when shorting is allowed but there is a cost: So far in this question we have made no assumptions about preferences beyond risk aversion. If one asset is considerably more volatile than the other then variance can be reduced by shorting the more volatile one. However (assuming positive mean μ) this will reduce the expected return. There will be a risk return trade off and a multi-valued efficient frontier can appear. However, calculating an "optimal" portfolio in this situation will require making assumptions about preferences over risk and return. For example if $\frac{(\sigma_Y^2 - \sigma_X\sigma_Y\rho_{XY})}{VAR(X-Y)}$ is indeed negative then w_{opt} only takes this value if the investor is *completely* agnostic to return - This is an implausible assumption.

Question 2

We use matrix notation and drop the t subscript (we need consider only contemporaneous covariances as the disturbances are iid across time). I write the variance of ϵ_i as σ_ϵ^2. There are n assets, indexed by i in the question and the variables have the following dimensions:

R	$n \times 1$ vector of (asset) returns.
R_m	Scalar market returns.
α	$n \times 1$ vector of constants.
β	$n \times 1$ vector of asset "market betas"; the sensitivity of the assets to the "market".
ϵ	$n \times 1$ vector of shocks. These represent the idiosyncratic variation to asset returns.
σ_m^2	Scalar variance of R_m
D	$n \times n$ diagonal covariance matrix of ϵ, equal to $\sigma_\epsilon^2 I_n$ as ϵ_i are iid

In the derivation we use the fact that the covariance of vectors is a bi-linear function and that $VAR(AB) = AVAR(B)A'$

$$
\begin{aligned}
\Omega &= VAR(\alpha + \beta R_m + \epsilon) \\
&= COV(\beta R_m + \epsilon, \beta R_m + \epsilon) && \because \alpha \text{ is a constant vector} \\
&= VAR(\beta R_m) + COV(\epsilon, \beta R_m) + COV(\beta R_m, \epsilon) + VAR(\epsilon) \\
&= \beta VAR(R_m)\beta' + VAR(\epsilon) && \because \epsilon \text{ and } R_m \text{ are iid} \\
&= \beta \sigma_m^2 \beta' + D \\
&= \sigma_m^2 \beta\beta' + D && \text{as required}
\end{aligned}
$$

Inspecting the matrix a little further gives some insight into the one factor market model. For $i \neq j$

$$
\begin{aligned}
COV(R_i, R_j) &= \Omega_{ij} = \sigma_m^2 \beta_i \beta_j \\
VAR(R_i) &= \Omega_{ii} = \sigma_m^2 \beta_i^2 + \sigma_\epsilon^2 \\
\rho_{ij} &= \frac{\Omega_{ij}}{\sqrt{\Omega_{ii}\Omega_{ij}}} = \frac{1}{\sqrt{\left(1 + \left[\frac{\sigma_\epsilon}{\sigma_m \beta_i}\right]\right)\left(1 + \left[\frac{\sigma_\epsilon}{\sigma_m \beta_j}\right]\right)}}
\end{aligned}
$$

The covariance of stocks increases with their relative sensitivity to the market (β). Variance of stock returns also increases β as well as the variances of the market (σ_m) and idiosyncratic shocks (σ_ϵ). The *correlation* of stocks increases with their β's and decreases with the ratio of the idiosyncratic variance to market variance (σ_ϵ/σ_m). This makes intuitive sense as the factor driving the correlation is the variation in the single market factor. Stocks whose returns are driven more by the market component will be more correlated and the overall level of correlation decreases with how much stocks move idiosyncratically compared to the market. In the extreme case of no idiosynchratic stock movements ($\sigma_\epsilon = 0$) returns of all stocks are perfectly correlated.

Question 3

This question, although concerning empirical finance, is really a question about OLS under model mis-specification. You should have come across these problems during the Econometrics course.

We assume that the market model holds and thus the true model is:

$$
\begin{aligned}
R_{it} &= \alpha_i + \beta_i R_{mt} + \epsilon_{it} \\
E(\epsilon_{it}|R_{mt}) &= 0
\end{aligned}
$$

So

$$
\begin{aligned}
Z_{it} &= R_{it} - R_{ft} = \alpha_i + \beta_i R_{mt} + \epsilon_{it} - R_{ft} \\
Z_{it} &= \alpha_i + \beta_i (R_{mt} - R_{ft}) + (\beta_i - 1) R_{ft} + \epsilon_{it}
\end{aligned}
$$

We can also write the model in a way so that the error term has zero expectation

$$
\begin{aligned}
Z_{it} &= (\alpha_i + [\beta - 1]E(R_{ft})) + \beta_i Z_{mt} + \eta_{it} \\
\eta_{it} &= [\beta - 1]\left(R_{ft} - E(R_{ft})\right) + \epsilon_{it} \\
E(\eta_{it}) &= 0
\end{aligned}
$$

The following regression is performed:

$$Z_{it} = \gamma_i + \delta_i Z_{mt} + \eta_{it}$$

Even if all Gauss-Markov assumptions hold (which they do not) $\hat{\gamma_i}$ will be biased for α_i as the intercept in the model is shifted by $[\beta - 1]E(R_{ft})$. We can also show the slope estimator is a biased for β using the OLS standard formula in terms of sample (co)variances:

$$\begin{aligned}
\hat{\delta}_i = \frac{\widehat{COV}(Z_{it}, Z_{mt})}{\widehat{VAR}(Z_{mt})} &= \frac{\widehat{COV}(\alpha_i + \beta_i Z_{mt} + (\beta_i - 1)R_{ft} + \epsilon_{it}, Z_{mt})}{\widehat{VAR}(Z_{mt})} \\
&= \beta_i + (\beta_i - 1)\frac{\widehat{COV}(R_{ft}, Z_{mt})}{\widehat{VAR}(Z_{mt})} \\
&= \beta_i + (\beta_i - 1)\frac{\widehat{COV}(R_{ft}, R_{mt} - R_{ft})}{\widehat{VAR}(Z_{mt})} \\
&= \beta_i + (1 - \beta_i)\frac{\widehat{VAR}(R_{ft})}{\widehat{VAR}(Z_{mt})} + (\beta_i - 1)\frac{\widehat{COV}(R_{ft}, R_{mt})}{\widehat{VAR}(Z_{mt})}
\end{aligned}$$

so

$$\hat{\delta}_i \longrightarrow \beta_i + (1 - \beta_i)\frac{VAR(R_{ft})}{VAR(Z_{mt})} + (\beta_i - 1)\frac{COV(R_{ft}, R_{mt})}{VAR(Z_{mt})} \neq \beta_i$$

Similarly, we can also show the Gauss-Markov assumption $E(\eta_{it}|Z_{mt}) \neq 0$:

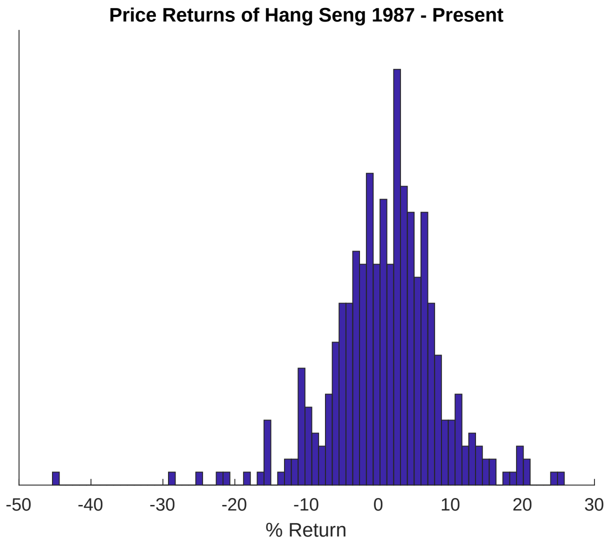
$$\begin{aligned}
E(\eta_{it}|Z_{mt}) &= E((\beta - 1)[\beta - 1]\left(R_{ft} - E(R_{ft})\right) + \epsilon_{it}|R_{mt} - R_{ft}) \\
&= (\beta - 1)E\left[\left(R_{ft} - E(R_{ft})\right)|R_{mt} - R_{ft}\right] + E(\epsilon_{it}|R_{mt} - R_{ft}) \\
&= (\beta - 1)E\left[R_{ft} - E(R_{ft})|R_{mt} - R_{ft}\right] \\
&\neq 0
\end{aligned}$$

Question 4

As we saw in the preceding class, stock returns are not observed to be normal. The distributions typically have large excess kurtosis and negative skewness. Fama has argued that using aggregation of daily returns to form monthly returns suppresses the higher moments (by the number of units of aggregation) which makes the distributions closer to normal. This has the downside that the sample size is reduced, and the argument is not beyond criticism.[1] Below I show the Hang Seng from 1987 along with sample statistics. The excess kurtosis has reduced from the daily returns but it is still high and the null of normality is rejected by the JB test.

The CAPM in the presence of a riskless asset (Sharpe-Lintner version) implies that excess returns of stocks are fully explained by their "β". This implies the intercepts (α) in the regression of excess stock returns on excess market returns is zero. It also implies the risk premia relationship $\pi_i = \beta_i \pi_m$. The former is tested by Maximum Likelihood equivalent to running equation-by-equation regressions

[1] For example, daily returns are also the aggregation of large numbers of intra-day returns but do not appear to be normal. The aggregation argument requires that these higher moments exist - it is not even clear that stock returns are well described by distributions with finite variance, on which the CLT relies; for example Mandlebrot in his 1963 paper "The Variation of Certain Speculative Prices" proposes the use of a family of distributions that have infinite variance.



Price Returns Statistics:
Mean 0.94
StdDev 7.73
Kurt 7.18
Skew -0.7
JBstat 305

of excess stock returns on excess market returns and the latter is tested by cross sectional regressions of the observed risk premia relationships. The fat tailed and skewed distributions of returns results in the fact that any tests of CAPM which rely on the normality assumption will not be valid. Thus only approaches that use large sample asymptotic results are robust (and even then there is the questionable assumption of finite variance of returns). In the context of the tests mentioned in lecture notes the LR and Wald tests based on the χ^2 distribution are valid whereas the exact finite-sample variant that uses the F statistic is not. Note that adjustments are required in the presence of heteroskedasticity and serial correlation of errors.

Question 5

We are asked to find the "statistical properties" of RV_n. Faced with this question one should probably calculate the first two moments. One could also consider RV_n as an estmator for the variance of the underlying process σ^2 and show it is consistent. This can be achieved in a fairly straight-forward manner using the the Strong LLN in the equally spaced case, and the Weak LLN in the unequally spaced case. In the latter, one must consider, say, $Y_i = r_i^2 - E(r_i^2)$. Then Y_i are independent and $E(Y_i) = 0 \, \forall \, i$ so WLLN can apply. Consistency trivially follows from the below derivations of the first two moments, as the expectation converges to σ^2 and the variance vanishses in the limit.

The complication in this question comes from the fact that the intra-day returns are not equally spaced. This means that although r_i^2 is independent they are not independent and identically distributed.

First lets work in the simpler equally spaced case. Say there are n trades then $t_i - t_{i-1} = \frac{1}{n}$ and

$$r_i \sim N(\frac{\mu}{n}, \frac{\sigma^2}{n}) \, \forall i$$

The CLT can be applied to the realised volatility. First calculate the mean and variance of r_i^2 (using $E(X^4) = \mu^4 + 6\mu^2 + 3\sigma^4$ when $X \sim N(\mu, \sigma^2)$).

$$\begin{aligned}
E(r_i^2) &= VAR(r_i) + E(r_i)^2 \\
&= \frac{\mu^2 + n\sigma^2}{n^2} \\
&= \frac{\sigma^2}{n} + O(\frac{1}{n^2}) \\
VAR(r_i^2) &= E(r_i^4) - E(r_i^2)^2 \\
&= \frac{\mu^4}{n^4} + 6\frac{\mu^2}{n^2}.\frac{\sigma^2}{n} + 3\frac{\sigma^4}{n^2} - [\frac{\sigma^2}{n} + O(\frac{1}{n^2})]^2 \qquad \text{using the expression for } E(X^4) \text{ above} \\
&= \frac{\mu^4 + 6n\mu^2\sigma^2 + 3n^2\sigma^4}{n^4} - \frac{\sigma^4}{n^2} + O(\frac{1}{n^3}) \\
&= 2\frac{\sigma^4}{n^2} + O(\frac{1}{n^3})
\end{aligned}$$

Then by the CLT

$$RV_n = \sum_{i=1}^{n} r_i^2 \Longrightarrow N(\sigma^2, 2\frac{\sigma^4}{n})$$

In fact, the exact distribution of RV_n for finite n in the equally spaced trade case is known. The realised volatility is the i.i.d. sum of non-central Normal Distributions squared which is in fact a non-central chi-squared distribution. This is beyond the scope of this course but those interested may wish to look it up on Wikipedia. When $\mu = 0$ the distribution is of course a standard chi-sq distribution, which should be apparent to you.

Turning to the more difficult case. As the r_i^2 are no longer i.i.d. a CLT cannot be applied (there is a version of the CLT that can be applied in the central case when $\mu = 0$ but this again is beyond the scope of this course). We calculate the mean and variance of RV_n directly. Writing $\tau_i = t_i - t_{i-1} \Rightarrow r_i \sim N(\tau_i\mu, \tau_i\sigma^2)$. Noting that $\tau_i = O(\frac{1}{n})$ (because n of them sum to unity):

$$\begin{aligned}
E(RV_n) &= E(\sum_{i=1}^{n} r_i^2) \\
&= \sum_{i=1}^{n} E(r_i^2) \\
&= \sum_{i=1}^{n} (VAR(r_i) + E(r_i)^2) \\
&= \sum_{i=1}^{n} (\tau_i\sigma^2 + \tau_i^2\mu^2) \\
&= \sigma^2 \sum_{i=1}^{n} \tau_i + \mu^2 \sum_{i=1}^{n} \tau_i^2 \\
&= \sigma^2 + \mu^2 \sum_{i=1}^{n} \tau_i^2 \qquad \because \sum_{i=1}^{n} \tau_i = 1 \\
&= \sigma^2 + O(\frac{1}{n}) \qquad \because \tau_i = O(\frac{1}{n})
\end{aligned}$$

the mean remains unchanged from the equally spaced case. Turning to variance and using $E(r_i^2) = \tau_i\sigma^2 + \tau_i^2\mu^2$ and $E(r_i^2) = \tau_i^4\mu^4 + 6\tau_i^3\mu^2\sigma^2 + 3\tau_i^2\sigma^4$:

$$
\begin{aligned}
VAR(RV_n) &= VAR(\sum_{i=1}^{n} r_i^2) \\
&= \sum_{i=1}^{n} VAR(r_i^2) \qquad\qquad \because r_i^2 \text{ are independent} \\
&= \sum_{i=1}^{n} \left[E(r_i^4) - E(r_i^2)^2\right] \\
&= \sum_{i=1}^{n} \left[(\tau_i^4\mu^4 + 6\tau_i^3\mu^2\sigma^2 + 3\tau_i^2\sigma^4) - (\tau_i\sigma^2 + \tau_i^2\mu^2)^2\right] \\
&= \sum_{i=1}^{n} \left[3\tau_i^2\sigma^4 - \tau_i^2\sigma^4 + O(\frac{1}{n^3})\right] \\
&= 2\sigma^4 \left(\sum_{i=1}^{n} \tau_i^2\right) + O(\frac{1}{n^2})
\end{aligned}
$$

so the variance of RV_n is still $O(\frac{1}{n})$ but is different than in the equally spaced case.

Higher moments of RV_m can be calculated in a similar fashion using the standard moment expressions for $N(\mu, \sigma^2)$. As always, look at Wikipedia for more details.