BUFN402: Portfolio Management

James Zhang*

March 26, 2024

These are my ("live-TEX"-ed) notes for UMD's BUFN402: Portfolio Management, which is a required class in the Computational Finance Minor. This course is taught by Professor Seokwoo Lee.

Contents

*Email: jzhang72@terpmail.umd.edu

§1 Intro

Hey guys, let's learn the skills needed to becomne junior quant researchers!

§2 Probability Theory

See Professor Lee's comprehensive slideshow on the probability theory concepts that we need to know.

§3 Simple PF Theory

Example 3.1

No Arbitrage Principle. Consider a **default-free zero-coupon bond** paying you \$1000 in one year. Assume risk-free interest rate is \$5 and the financial market is well functioning and competitive. Suppose the bond is traded for \$940. Is there an arbitrage opportunity?

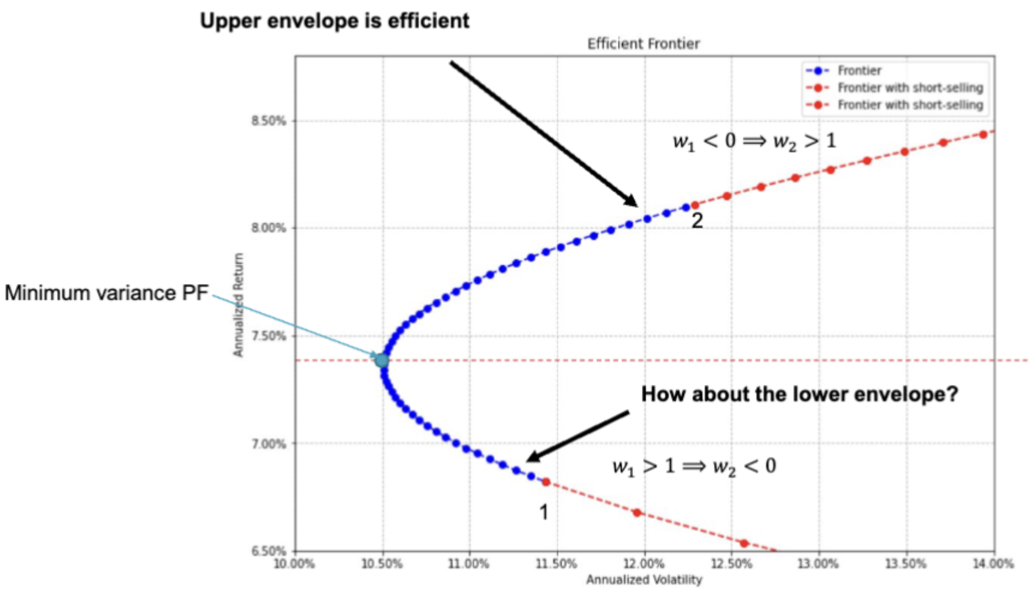
Solution. Buy the bond for \$940. At the same time, borrow $1.05x = 1000 \implies x = 952.38$ from the bank. In 1 year, we get 1000 from the bond which we give to the bank. Therefore, we earn 12.38 in cash today, without taking any risk or paying any of our money in the future. □

What if the price of the bond is 960. Is there an arbitrage opportunity?

Solution. Sell the bond for 960, then invest 952.38 at the bank. In a year, it grows to 1000 and then you use it to pay off the bond. You keep 7.62. □

Definition 3.2. In financial markets, it is possible to sell a security you do not own. This is called **short-selling**. In short-selling, an investors borrows the security, sells it, and then must either return the security by buying it back or pay the owner the cash flows he would have received.

Note 3.3. The example above assumed no transaction costs and funding costs.



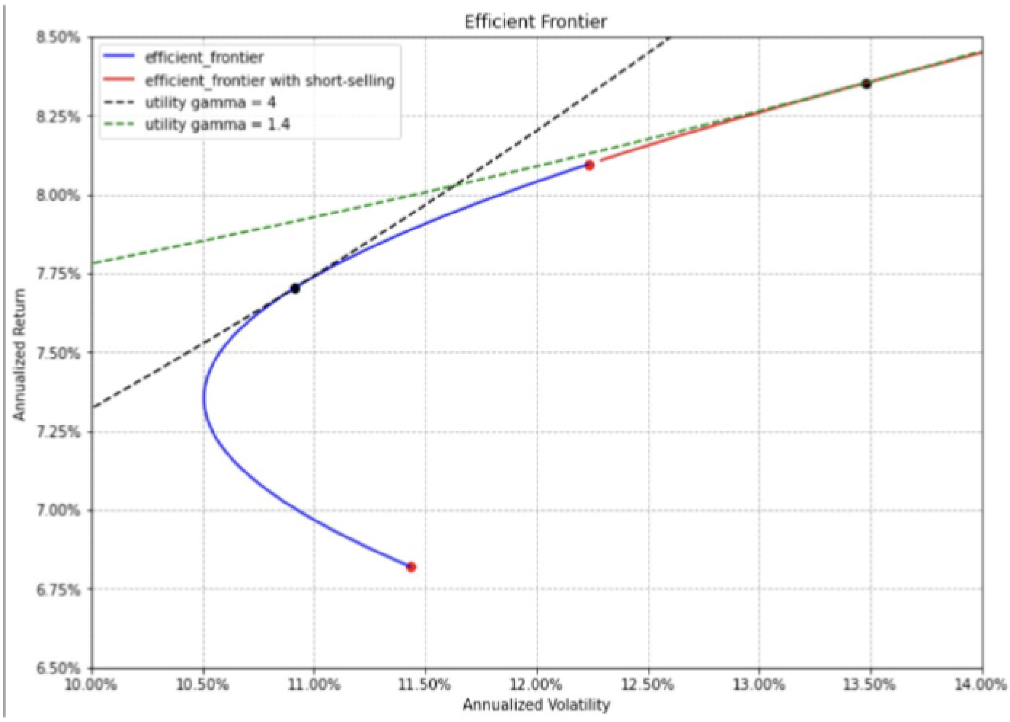
Definition 3.4. The **efficient frontier** defines the best feasible set of PFs, given two asset characteristics (mean and variance). Note that above, the red portions come from allowing short selling. Also note the minimum variance PF. Additionally, note that all PFs below the dotted red line are **inefficient** - that is, for the same level of risk, you can obtain a PF with greater expected return.

Definition 3.5. A **mean-variance utility maximizer** wants to solve

$$\max_{w_1} \mathbb{E}(R_p) - \frac{\gamma}{2}\text{Var}(R_p)$$

where γ is known as **risk-aversion** and sometimes $\tau = \frac{1}{\gamma}$ is **risk-tolerance**.

§3.1 Frontiers with Risk-Free Assets



Remark 3.6. In addition to 2 risky assets, suppose the investor can freely borrow and lend at r_f. Consider a PF combining any PF on the frontier with the risk-free asset. Then all feasible efficient PFs become on a line.

Solution. Let weight on a risky PF R_p on the frontier be w_p and so $(1 - w_p)$ to the risk-free asset. The return of the PF is then

$$R_{pf} = w_p R_p + (1 - w_p) r_f$$

$$\mathbb{E}(R_{pf}) = w_p \mu_p + (1 - w_p) r_f$$

$$\sigma(R_{pf}) \;= w_p \sigma_p \implies (\sigma(R_{pf}, \mathbb{E}(R_{pf})))$$

is linear in w_p where the slope of this line is $\frac{(u_p - R_f)}{\sigma_p}$, **Sharpe Ratio**. □

Note 3.7. The optimal PF for the mean variance utility maximizer becomes where their utility function intersects with the line of maximum Sharpe Ratio.

§3.2 Long-Short Strategy of Hedge Fund

Take notes on this later. Optional for now.

§4 Capital Asset Pricing Model (CAPM)

Example 4.1

Let us construct an algorithm to find the efficient frontier with N risky assets. Vary target expected returns of the PF. Given μ_p, minimize the volatility of the PF by finding the weights subject to

- $\sum w_i = 1$
- $\mathbb{E}(R_p) = \sum w_i \mathbb{E}(R_i)$
- $0 \leq w_i \; \forall \; i \in [1, N]$

Note that the last constraint is omitted if short selling is allowed. Additional constraints can be added to customize the PF selection process.
More formally, the optimization step is states as follows. For a given target expected return r_p:

$$\min_w w^T \sigma w \text{ subject to } 1^T w = 1, w^T \mu = \mu_p, w \geq 0$$

To find the **tangent PF**, which ist he PF with maximum Sharpe Ratio, solve

$$\max_w \frac{\mu^t w - r_f}{\sqrt{w^T \sigma w}} \text{ subject to } 1^T w = 1, w \geq 0$$

The **global minimum variance PF** is easy to find in this methodology.

Note 4.2. Suppose now that your investment universe is all tradeable assets. Then the tangent PF is efficient and contains all assets.

Note 4.3. IMPORTANT: even if two investors have different risk tolerances, they still hold the **same risky tangent PF** T just with different weight to the tangent PF weights. Additionally, each weighting of the tangent PF is directly correlated with the market. This is left as an exercise, it involves computing the correlation coefficient between a weighted tangent PF and the market and showing that it equals 1.

§4.1 CAPM Assumptions and Consequences

Note 4.4. Let us make four key assumptions for CAPM.

a. Investors are mean-variance optimizers and take the market price as given.

b. Two-periods - investors are planning to hold for one identical period

c. Investors all face the same investment opportunities, characterized by N risky assets , and they all borrow and lend at the same equilibrium risk free rate.

d. **Rational Expectations** - all investors correctly interpret any info they have and all information contained in asset prices, so all investors will agree on expected returns, variances, and covariances of stocks.

Note 4.5. Some important consequences of the above:

a. Each investor agrees on all assets' positions in (μ, σ) space.

b. Each investor identifies the same PF as having the highest Sharpe - same tangentPF

c. All investors demand the same tangent PF. Recall that different risk tolerances just adjust the weight of the risk-free asset and the tangent PF

d. Because a security is owned by some, the sum of all investors' PFs must equal the market PF.

e. If a security in the market were not part of the efficient F, then no investor would want it, and there is an excess supply of this asset. This security's price would call, causing expected return to rise, until it became an attractive investment. In this way, prices in the market will adjust so that the effiicent PF and the market PF coincide, and demand equals supply. This is essentially just a proof by contradiction.

f. CAPM theory identifies the efficient PF as the market PF, and the market PF is easier to identify (SP500)

g. Thus, if a mean-variance investor just performs regressions and not necessarily the mean-variance analysis themselves, the investor can just use the market PF as the optimal tangent efficient PF.

§4.2 CAPM Derivation

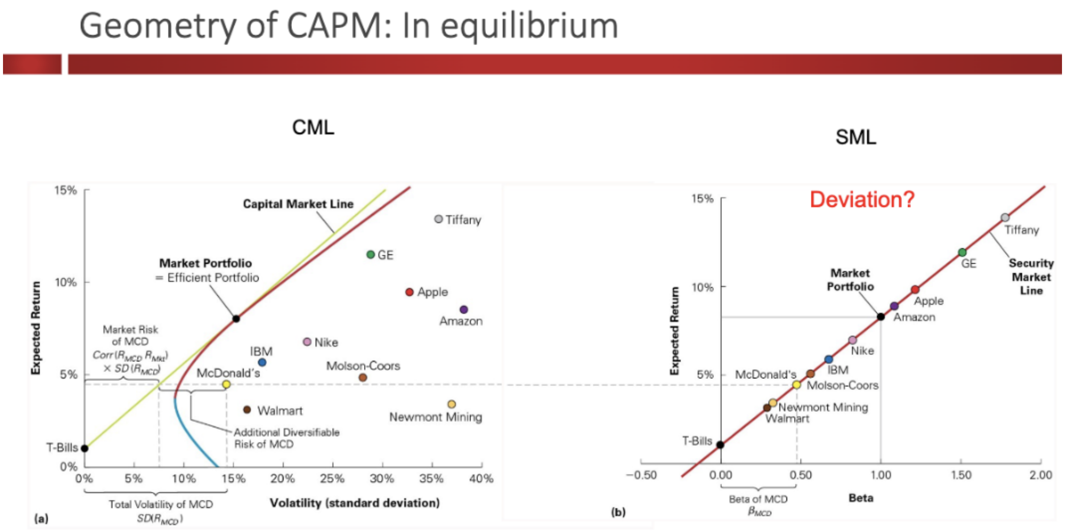
Proof.

$$\vdots$$

Therefore, the market PF is indeed mean-variance efficient. Hence

$$\mathbb{E}(R_i) - r_f = \beta_{i,m}(\mathbb{E}(R_m) - r_f)$$

Excess returns of the asset can be perfectly explained by the excess returns of the market times some scalar, β. □



Note 4.6. Note that in the **Security Market Line** graph above, all assets fall on the same line, whose slope is $\mathbb{E}(R_M) - r_f$ which is the excess returns of the market.

Note 4.7. Theoretically, if CAPM were true, one could arbitrage when an asset fell off the **SML**.

§5 Inference and Regressions

§5.1 Inference

Definition 5.1. An **estimator** is a function of a sample data frawn randomly from a population. For example, the sample mean is an unbiased estimate of the population mean.

Note 5.2. An estimate is a random variable!

Definition 5.3. $\bar{X}_n$ is the best estimator of μ because $\bar{X}_n$ is **unbiased** meaning the expected value is the true value, and $\bar{X}$ is **consistent** meaning

$$\lim_{n\to\infty} \bar{X}_n = \mu$$

Definition 5.4. Since $\bar{X}_n$ or any estimator is a random variable, its properties are determined by its distribution. We call the distribution of $\bar{X}_n$ as the **sampling distribution** of the estimator.

Definition 5.5. **IID = independent and identically distributed**

Definition 5.6. An **effiicent** estimator is the estimator with the smallest variance amongst all unbiased estimators of the true value.

Example 5.7

For a fixed (large) n, we have

$$\mathbb{P}(-z_{\alpha/2} \leq \frac{\sqrt{n}}{\sigma}(\bar{X}_n - \mu) \leq z_{\alpha/2}) = 1 - \alpha$$

which is equivalent to

$$\mathbb{P}(\bar{X}_n - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \leq \mu \leq \bar{X}_n + \frac{\sigma}{\sqrt{n}} z_{\alpha/2} = 1 - \alpha)$$

and so the random interval

$$[\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2}, \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}]$$

contains the unknown mean μ with probability $1 - \alpha$. We call this interval estiamte with critical value α as the $100(1-\alpha)\%$ **confidence interval** for μ.

Example 5.8

The above assumed that we know the true variance σ^2. If we have unknown variance, then we use the unbiased estiamte

$$\sigma_n^2 = \frac{1}{n-1} \sum_i (X_i - \bar{X}_n)^2$$

Then obtain the t-statistic and replace σ with s_n.

Note 5.9. When $n >> 120, t_{n-1}$ is close to a standard normal distribution $N(0, 1)$.

$$\frac{(\bar{X}_n - \mu)}{s_n/\sqrt{n}} \sim t_{n-1}$$

Note 5.10. Statisticians often refer to the standard deviation of $\bar{X}_n$ as **standard error** of the estimator

$$se(\bar{X}_n) := \text{estimator}(\sqrt{\text{Var}(\bar{X}_n)}) = \frac{(\bar{X}_n - \mu)}{se(\bar{X}_n)} \sim t_{n-1}$$

$$se(\bar{X}_n) = \frac{s_n}{\sqrt{n}}$$

Theorem 5.11

Central Limit Theorem: Suppose that $(X_1, \cdots X_n)$ are iid of any distribution with $\mathbb{E}(X_i) = \mu$ and $\text{Var}(X_i) = \sigma_2 < \infty$ then as $n \to \infty$

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \to N(0, 1)$$

but there are limitations which we will get to in the regression lectures.

§5.2 Regression

Definition 5.12. The simple definition of a regression is

$$y = \beta_0 + \beta_1 x + \epsilon$$

where y is the dependent variable, β_0 is the intercept, β_1 is the slope parameter, x is the independent variable, and $\epsilon \sim N(0, \sigma^2)$ is noise

Definition 5.13. OLS says that the best values of $(\hat{\beta}_0, \hat{\beta}_1)$ should minimize residuals

$$\min_{(\hat{\beta}_1, \hat{\beta}_2)} (y_i - \hat{y}_i)^2$$

$$\hat{\beta}_1 = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} \approx \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum_j (x_i - \bar{x})^2} \text{ and } \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Example 5.14

We know that OLS estimators are random variables, so we can study their distributions. Recall that the error term $\epsilon \sim N(0, \sigma^2)$, and so the OLS estimators $\hat{\beta}_0, \hat{\beta}_1$ are also normally distributed. Recall

$$\hat{\beta}_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum_j (x_i - \bar{x})^2}$$

which is a weighted sum of the y_i by keeping x_i fixed. Under this, consider

$$y_i = \beta_0 + \beta_1 x_i + \epsilon$$

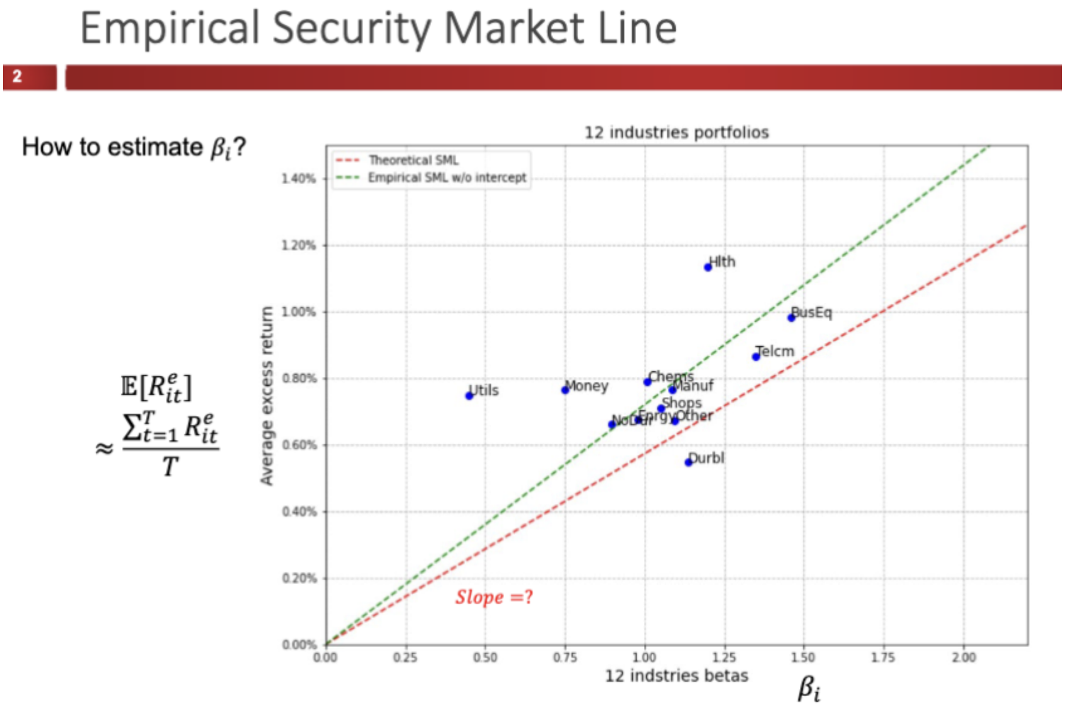
and so y_i is normal because $\epsilon \sim N(0, \sigma^2)$ and so $\hat{\beta}_1$ is also normal. Similarly, $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1$ is also a linear combination of y_i.

Note 5.15. Note that under some regular conditions, then the OLS estimators $\hat{\beta} := (\hat{\beta}_1, \hat{\beta}_0)$ are unbiased and consistent.

More on distributons of OLS estimators with unknown σ, confidence intervals for OLS estimators, hypothesis testing of the slope β_1, and the Gauss-Markov Theoerem.

§6 CAPM Anomalies

The CAPM is not particularly good at explaining asset returns, consider the following



Note 6.1. The CAPM model is

$$\mathbb{E}(R_{it}) - r_f = \alpha_i + \beta_{im}(\mathbb{E}(R_{mt}) - r_f)$$

where the main test is $\alpha_i = 0 \ \forall \ i$. The main model is a cross sectional regression but β_i are estimated as **time-series regressions**.

Note 6.2. The CAPM states that the expected return of the stock should only depend on β, the amount of systematic risk of the stock and the fraction of total risk correlated with the market PF.
Nothing else other than a firms β should predict return. Note that in the CAPM, high β (systematic risk) should present higher returns. Think economics: higher risk, higher returns.

Definition 6.3. Recall that **alpha** α_i is any residual return of stock i that cannot be explained by β_i. It is an abnormal return after adjusting for systematic risk of the stock i.

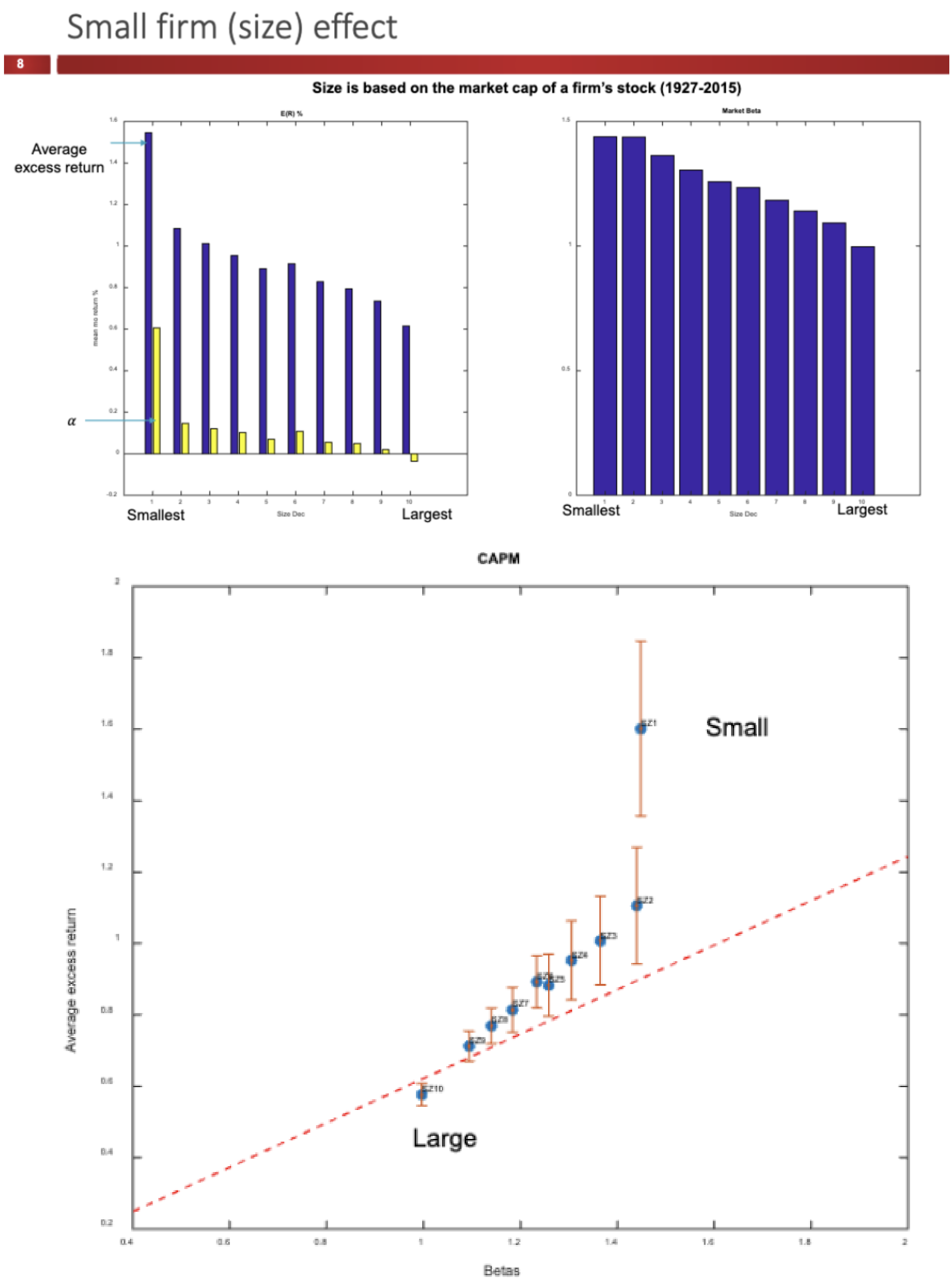
Remark 6.4. The two main anomalies with respect to the CAPM are

1. **Positive alpha**: some stocks still present positive returns after adjusting thier systematic risk, which means something else than β can predict returns!

2. **Low beta yields higher returns high beta**, which contradicts economics...

§6.1 Anomaly 1: Small Firm Effect

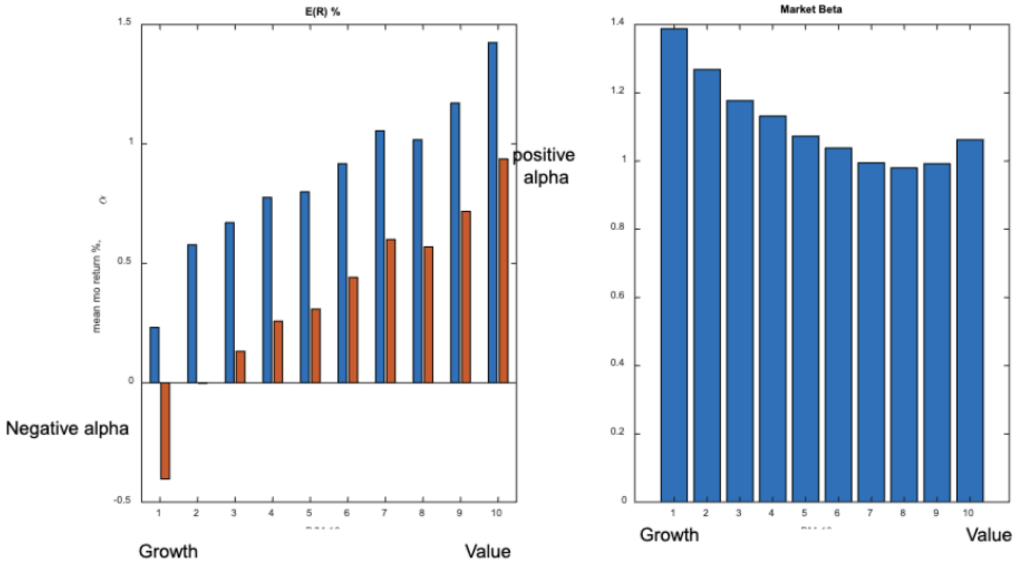
Definition 6.5. Stocks of small firms have earned more risk-adjusted returns than large firms.

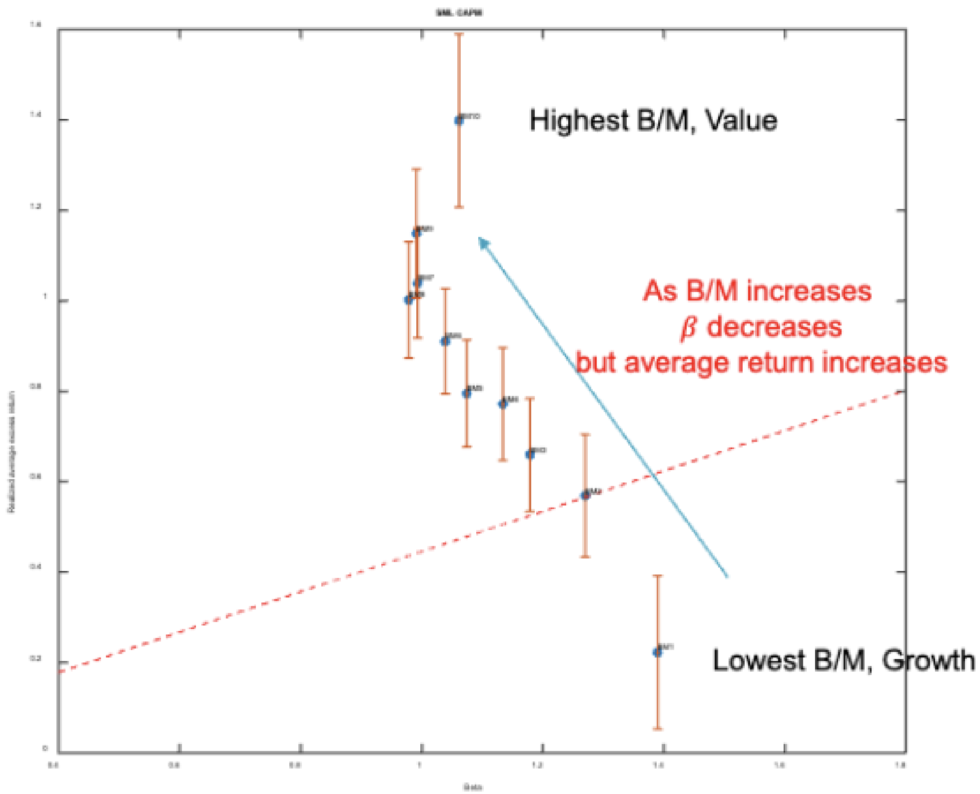


Note that almost all PFs earn positive alpha, which is the anomaly. Betas are not an anomaly here, since bigger betas earn higher returns.

§6.2 Anomaly 2: Book to Market Effect

Definition 6.6. Stocks with high B/M (value stocks) tend to outperform stocks with low B/M ratio (growth stocks).

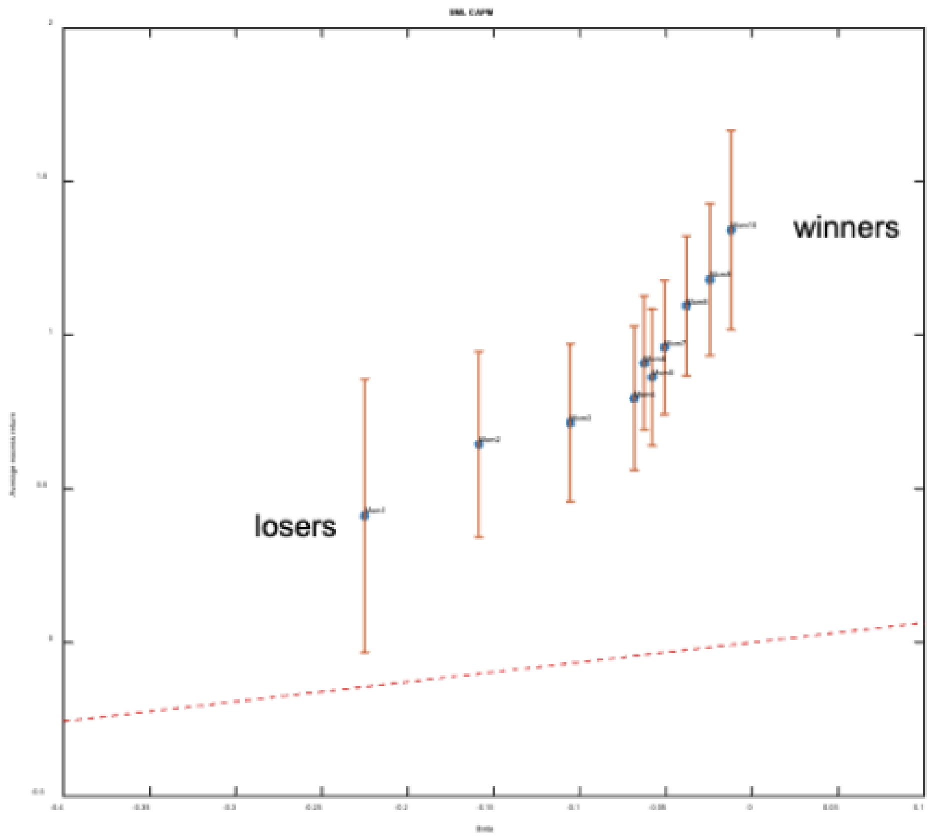




There are two anomalies here: betas move the opposite direction, and majority of PFs earn a substantial positive alpha.

§6.3 Anomaly 3: Momentum Effect

Definition 6.7. Stocks that have done well (or poorly) in the previous year tend to do well (or poorly) in the future.



This goes against the CAPM since past returns (other than beta) should not predict the returns.

Note 6.8. It is likely that the CAPM fails because of behavioral biases. Some uninformed traders are systematically attracted to growth stocks that receive greater news coverage. Thus, they sell winners and hang on to losers.

§7 Multivariate Regressions

§7.1 Statistics

Definition 7.1. A multiple linear regression model captures

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \epsilon$$

The motivation for this model is to incorporate more explanatory factors into the model.

Note 7.2. The method of estimating β's is the same as univariate linear regression, just minimize the sum of squared residuals.

Note 7.3. Now intead of thinking of betas as whole derivatives, they are really

$$\beta_j = \frac{\partial y}{\partial x_j}$$

and they betas are also conditioned on the other factors in the model.

Example 7.4

Suppose you run a linear regression and get

$$y = 0.6x_1, \hat{\beta}_1 = 0.6$$

Then you incorporate a new factor x_2 and get

$$y = 0.6x_1 + 0.2x_2, \hat{\beta}_1 = 0.6, \hat{\beta}_2 = 0.2$$

then x_1, x_2 must be independent by $\mathbb{P}(A|B) = \mathbb{P}(A) \implies$ independence.

Example 7.5

Suppose we regress average test scores on student teacher ratio and percnet of English as Second Language. In the univariate regression, we get

$$y = 698 - 2 * STR$$

and in the multivariate regression we get

$$y = 686 - 1.1STR - 0.65ELpct$$

This difference in the STR coefficients occurs because STR is correlated to $ELpct$.

More on goodness of fit, adjusted r-squared, hypothesis testing for β_i, testing the mdoel as a whole using F-statistic, multicollinearity

§7.2 Multifactor Models

Note 7.6. The major difficulty with single factor analysis so far is that there are many other sources of common variiation in stock returns than just the market return, such as cyclical/seasonal stocks or firms in the same industry generally also tend to move together.

Definition 7.7. Generalize our analysis from before into a **multifactor model**. If there are K factor portfolios (THESE ARE NOT ASSET SPECIFIC) capturing common influence of K underlying sources of risk, then we have

$$R_{it} - r_f = \alpha_i + \sum_{k=1}^{K} \beta_{ik} F_k + \epsilon_{it}$$

Once more, assume that residuals are uncorrelated across stocks.

Note 7.8. There should be an economic story behind each of the K factors for why they are being included in the analysis, although statistical models are fine.

§7.3 Fama-French Factor Models

Definition 7.9. In 1993, Fama and French published the **Fama-French 3 Factor Model (FF3)**

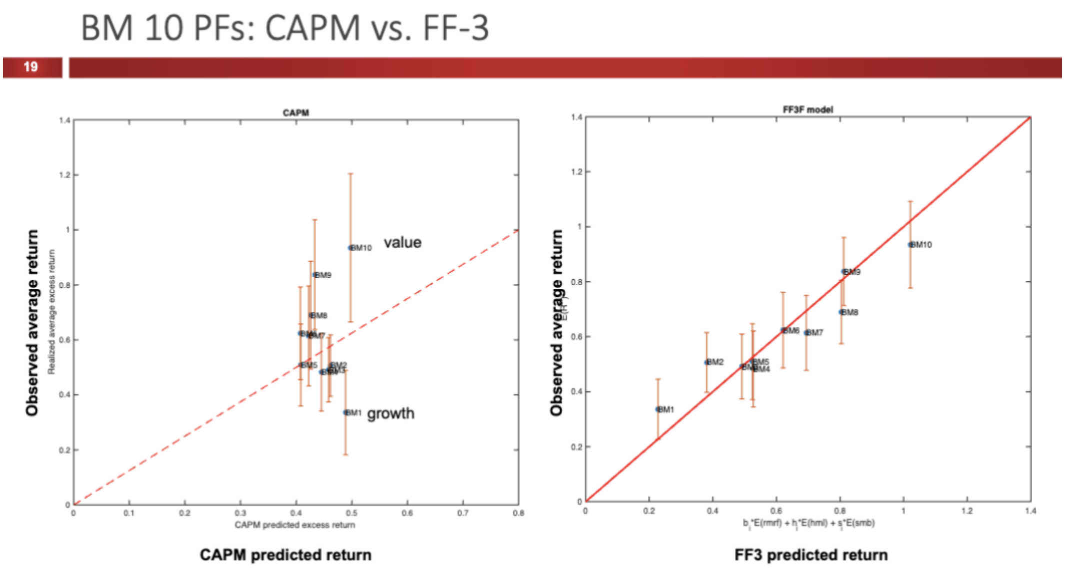
$$\mathbb{E}(R_i) - r_f = \beta_{i,MKT}(\mathbb{E}(R_m) - r_f) + \beta_{i,SMB}\mathbb{E}(R_{SMB}) + \beta_{i,HML}\mathbb{E}(R_{HML})$$

The associated multifactor regression equation is

$$\mathbb{E}(R_i) - r_f = \alpha_i + \beta_{i,MKT}(\mathbb{E}(R_m) - r_f) + \beta_{i,SMB}\mathbb{E}(R_{SMB}) + \beta_{i,HML}\mathbb{E}(R_{HML}) + \epsilon_{it}$$

where

- MKT: the same as the CAPM, the market-risk premium
- SMB: return on the benchmark PF of small firms minus returns on big firms
- HML: return on benchmark PF of value firms minus return on growth firms
- $\beta_{i,x}$ is the overall sensitivity to the x factor and correlation between return on the asset and the x strategy returns



Note 7.10. Researchers have tried additional factors such as momentum and liquidity.

§7.4 How to make a factor?

Note 7.11. The typical procedure is

1. Sort stocks on a characteristic into a quintile or decile portfolios and document a pattern in average returns
2. Construct a long-short strategy that buys the top and shorts the bottom
3. This difference in returns (factor) is then used as a factor to explain returns

By forming a long-short strategy, we "net out" some passive exposures.

Note 7.12. As an alternative to portfolio sorts, you can show that a characteristic predicts returns in the cross section using the **Fama-Macbeth procedure**. This is used to make **characteristic-based factors.**

Example 7.13

Consider an example where you run CAPM and you get $\alpha = 0.0061$. Using FF3, you get $\alpha = 0.005$. Therefore, controlling for size and value knocked about 1% off alpha, meaning adding the SMB and HML factors did something; they helped explain some of that unexplained returns. If SMB and HML did nothing, then market beta would stay the same.

- SMB loading is -0.5 means the strategy acts the otherway from small stocks, in fact you're overweighitng large stocks (large stock exposure)
- HML loading is 0.42 means you're overweighting value stocks.

The conclusion is that this strategy still yields considerable profit (positive alpha) even after controlling size and value effects.

§8 Practice Problems

#1. Suppose there are two risky assets in the economy, and traders can lend and borrow at the risk-free rate. Let the returns of the two risky assets are R_1 and R_2, the volatilities σ_1 and σ_2, and their covariances are σ_{12}. The traders can short-sell the risk assets.

a) State the optimization problem to find the tangent portfolio.

Solution.

$$\max_w \frac{\mu_p - r_f}{\sigma_p}$$

where $\mu_p = w_1\mu_1 + (1 - w_1)\mu_2$ and $\sigma_p = \sqrt{w_1^2\sigma_1^2 + (1 - w_1)^2\sigma_2^2 + 2w_1(1 - w_1)\sigma_{12}}$ □

b) Suppose $\mathbb{E}[R_1] = 0.1$ and $\mathbb{E}[R_2] = 0.2$. Assume that two risky assets are uncorrelated and each has variance equal to 1. There is a risk-free asset with return $r_f = 0.05$. What is the tangent portfolio?

Given the above, see that

$$\mu_p = w_1(0.1) + (1 - w_1)0.2 = 0.1w_1 + 0.2 - 0.2w_1 = 0.2 - 0.1w_1$$

$$\sigma_p = \sqrt{w_1^2 + (1 - w_1)^2}$$

Thus, our optimization problem becomes

$$\max_w \frac{0.2 - 0.1w_1 - 0.05}{\sqrt{w_1^2 + (1 - w_1^2)}} = \max_w \frac{0.15 - 0.1w_1}{\sqrt{w_1^2 + (1 - w_1^2)}}$$

$$(0.15 - 0.1w_1)(w_1^2 + (1 - w_1^2))^{-1/2}$$

The FOC with respect to w is (do the quotient rule rip)

$$\vdots$$

$$w_1 = 0.25 \implies (1 - w_1) = 0.75$$

#2. Using the data $\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\}$, you run a regression of y on x and get the estimated equation $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ where $\hat{\beta}_0$ *and* $\hat{\beta}_1$ are the usual Ordinary Least Square (OLS) estimators. Now, you run a regression of y on x with the duplicated data and get the estimated equation $\hat{y}_i = \tilde{\tilde{\beta}}_0 + \tilde{\tilde{\beta}}_1 x_i$. That is, you run the regression of y on x to fit the data $\{(x_1,y_1),(x_1,y_1),(x_2,y_2),(x_2,y_2)\cdots,(x_{n-1},y_{n-1}),(x_{n-1},y_{n-1}),(x_n,y_n),(x_n,y_n)\}$. Note that $\tilde{\tilde{\beta}}_0$ and $\tilde{\tilde{\beta}}_1$ are also the Ordinary Least Square estimators from the regression with this duplicated data.

a) Is $\hat{\beta}_1 = \tilde{\tilde{\beta}}_1$? If so, explain the reason why. If not, explain why not. For the full credit, be precise.

Proof. Yes, consider the scatter plot. The duplicate data has exactly the same points as the original data, so the slope estimates must be the same. Furthermore, consider the process of minimizng the same squared errors

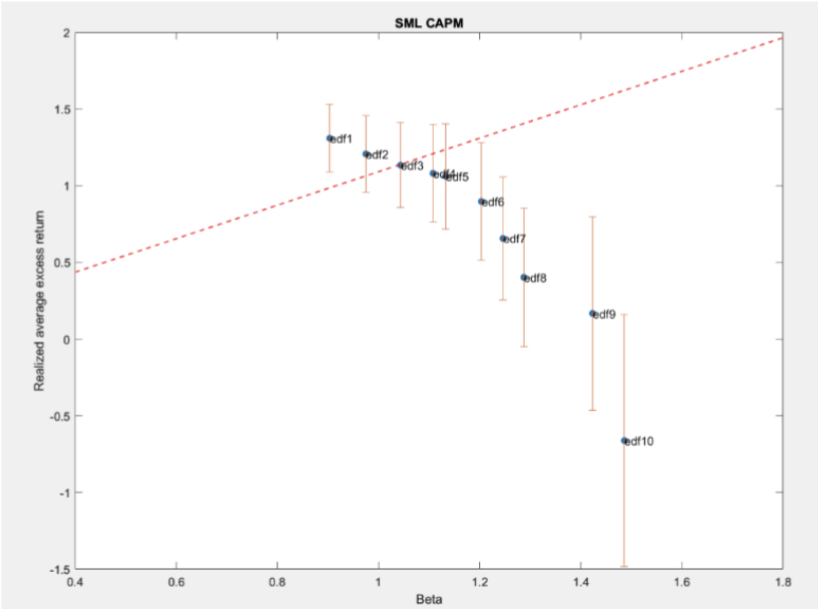
$$\min_{\beta_0,\beta_1} \sum_{i=1}^{2n} (y_i - \beta_0 - \beta_1 x_1)^2 = \min_{\beta_0,\beta_1} 2 \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_1)^2$$

will yield exactly the same OLS estimates $\hat{\beta}_0, \hat{\beta}_1$. □

b) Is the t-statistic of $\hat{\beta}_1$ same as the t-statistic of $\tilde{\tilde{\beta}}_1$? If so, explain the reason why. If not, explain why not.

Proof. □

[#3-#4]. Consider the following figure.



The horizontal axis of the figure represents the betas of the decile portfolios of expected default risk. The vertical axis depicts the realized returns of the portfolios. "edf 10" is the portfolio of the firms with the highest expected default risk whereas "edf 1" is the portfolio of the firms with lowest expected default risk. The red dashed line represents the (theoretical) security market line. The red solid bars represent the 95% confidence intervals of the realized returns.

Assume that the CAPM is correct model in this economy.

#3. Answer the following question:

a. Are there any anomalies with respect to the CAPM?
b. How can you make a trading strategy to exploits the anomalies, if any?

Solution.

a. Yes; many portfolios have nonzero alpha and as betas increase, decrease.

b. Long edf1 and short edf10

□

#1. Let's assume that we know the population is **normally** distributed with unknown mean μ and known variance σ^2. In other words, a sample $X_1, X_2 \cdots, X_n$ is independently drawn from an identical normal distribution with μ (unknown) and σ (known). What is the distribution of $\overline{X}_n = \frac{1}{N}\Sigma_{i=1}^{N} X_i$?

Note: we refer to this distribution as "sampling distribution of the sample mean."

Proof. Note that if each X_i is normal, then the sum of the X_i is normal.

$$\mathbb{E}(\frac{1}{N}\sum X_i) = \frac{1}{N}N\mu = \mu$$

$$\text{Var}(\frac{1}{N}\sum X_i) = \frac{1}{N^2}\text{Var}(\sum X_i) = \frac{1}{N^2}N\sigma^2 = \frac{\sigma^2}{N}$$

because the X_i are iid and so the covariances are all 0. Therefore,

$$\bar{X}_n \sim N(\mu, \frac{\sigma^2}{N})$$

□

#2. Suppose you obtained the sample average of hourly earnings is \$23.64 with the sample size $n = 100$. That is, $\bar{X}_{100} = \frac{1}{100}\sum_{i=1}^{100} X_i$. We also know that the population variance σ^2 = \$16. What is the 95% confidence interval of the population mean μ? Use the z-value $|z_{0.025|} \approx 1.96$.

Proof. Population variance is known, and so we can use the z-statistic and not the t-statistic. Our confidence is

$$[\bar{X}_{100} \pm \frac{4}{10} * 1.96]$$

□

#1. Show that when CAPM holds in equilibrium, all traders' optimal portfolio is perfectly correlated with the market portfolio, regardless of their risk-tolerance.

Proof. Note the important fact of the CAPM that despite varying risk-tolerances, all investors hold the tangent portfolio, just weighted differently towards the tangent portfolio and the risk free asset. Thus, we seek to show that

$$\rho = \frac{\text{Cov}(w * R_m, R_m)}{\sqrt{\text{Var}(w * R_m)}\sqrt{\text{Var}(R_m)}} = 1$$

$$= \frac{w\text{Var}(R_m)}{w\text{Var}(R_m)} = 1$$

Recall the property that $\text{Cov}(aX, Y) = a\text{Cov}(X, Y)$ for the numerator and $\text{Var}(aX) = a^2\text{Var}(X)$ in the denominator. □

#6. Consider the following single-variable regression:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

Assume that $\epsilon_i \sim^{iid} N(0,\sigma^2)$. Then, show that the estimator $\widehat{\beta_1}$ is also normally distributed, given the observed x_i.

Proof. Note that

$$\hat{\beta}_1 = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \approx \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_j (x_j - \bar{x})^2} = \sum_i c_i y_i$$

where

$$c_i := \frac{x_i - \bar{x}}{\sum_j (x_j - \bar{x})^2}$$

Fixing $x_i, y_i | x_i \sim N(\beta_1 x_i, \sigma^2)$ and so since $\hat{\beta}_1$ is the sum of normal distributions, it is also a normal distribution.

□