Jun Pan
junpan@mit.edu

MIT Sloan School of Management
15.433-15.4331 Financial Markets, Fall 2015

253-3083
E62-624

Class 1: Introduction

This Version: September 8, 2016

I'll write a note for each class. As you will see, these lecture notes are often more conversational, not as formal as a textbook. By allowing myself to be flexible and spontaneous in writing these notes, I actually get a lot of joy by just writing. Hopefully, they bring some joy to your reading as well. As of July 2016, I have been at MIT Sloan for 16 years. Teaching is not something that came naturally to me and I had my ups and downs. But one thing I know very well by now is that I enjoy teaching the most when I am sharing — knowledge, excitement, amazement, rational analysis, calm calculation, and quiet appreciation. So if I was not able to convey these fully in a classroom setting, I hope to use these lecture notes to get a second chance.

1 What to Expect from This Class?

- **This is not going to be a dry or dull class** because, as a subject matter, Finance is just too exciting to be taught in a boring way. Finance and financial markets are in real time and their ups and downs affect most people's lives.

  Last August, I was in Trout Lake, WA for a retreat. I thought I was as far away from Wall Street as possible. The location was rural America at its best and my fellow retreat participants are your typical spiritual type whose resume will not pass the first round of screening for any Wall Street jobs. Not that they are interested in such a job anyways. And yet, when the teacher mentioned in passing that the Dow had just lost 1,000 points that morning, there was a wave of agitation in that secluded woods we were in. "Everything was just fine before he had to mention that!" The woman sitting in front of me turned around and shared her thoughts during the break. I thought she was joking and was about to add mine when she murmured to herself, "Now I have to worry about my retirement money."

  In fact, this was not the first time when exciting financial news was broadcast to me in a retreat. When the markets went crazy in August 2011 because of the fear of contagion of the European debt crisis, I was in a meditation retreat in Serpentine, Australia. It was as far away from any known financial centers as you can imagine and we were not allowed to have Internet access. One day, I was mindfully doing my walking meditation, and the caretaker of the retreat center literally stopped me in my tracks and broadcast the news to me. It was only much later did I learn that he was doing currency speculation on the side and had just lost a lot of borrowed money that day.

For the famous week of Lehman bankruptcy (Monday) and AIG bailout (Tuesday), I was again in a meditation retreat. This time, I was not broadcast any news – this retreat center is known for its strictness. One week later, I was in Boston, digesting the shock. My conclusion: I should not be going to any retreat anymore. But more seriously, these examples bring home to us the widespread impact of the financial markets. It would be a pity to view Finance simply as formulas made up of Greek letters.

- **I'll teach this class as a professor, not as a professional investor.** Finance as an academic discipline has played and continues to play an important role in how Finance is practiced in real life. It builds theoretical models and frameworks, through which the seemingly random events in the financial markets can be analyzed, understood and quantified. It offers pricing models to facilitate the tremendous innovations in financial products. More often than not, the most creative ideas and the best trading strategies arise from research papers written by Finance professors. In the early days of the 70s, such Finance professors were called "academia nuts." Today, Finance professors are often heavily sought after and their research papers carefully studied and followed by practitioners. MIT Sloan is highly respected in Wall Street, not because of the professional investors we helped to produce, but because of the academic work done by Paul Samuelson, Franco Modigliani, Bob Merton, Fischer Black, Myron Scholes, and many other professors.

  Over the years, I enjoy many conversations with professional investors. Their anecdotes are often fun and exciting and one can easily spend an hour talking about a particular trade or event. Such anecdotes add more texture and color and will be used periodically in the class. But without a systematic framework, anecdotes remain just anecdotes. I know that many students are eager to get into the "real world" and shun the classroom materials as being too academic. Well, let me be honest here. Sooner or later, you are going to be in the "real world." Where else? And you will spend your days being buried in anecdotes and events. What is the hurry? Being in school is a rare opportunity in your life to retreat from the world and observe the events from a higher vantage point. Instead of being so eager to jump into the sea of uncertainty, why not first learn some basics about navigation?

- **This class will be an empirically driven class.** Over the years, Finance has become an empirically relevant discipline. Just imagine the amount of financial data that a Bloomberg terminal has to endure second by second, or millisecond by millisecond. When I was a PhD student in the late 1990s, I was given a quota of several gigabytes on a Sun Sparc workstation and I was thrilled. Today, that several gigabytes can hold maybe several minutes of stock trading data. How do we make sense of this jungle of data? Indeed, with the increasing availability of data and computing power, researchers have started to test the models developed in the 60s and 70s. By late 1990s, empirical work has replaced theoretical work as a more active research field in Finance. Studying the financial markets through the combination of data and theory is something I feel passionate about in my research. I find it exciting to be able to extract information from the seemingly noisy financial data. Analyzing and quantifying the regularity of financial risks from uncertain events is something I enjoy doing as a researcher. And I hope to be able to pass on this passion to you and help you develop these empirical skills, in addition to teaching you the basics of the financial markets. As such, this class will be an empirically driven class. We will talk theory, but theory is more of a guideline, like a map. This class is not about studying the map. It's about walking the path.

- **All class materials are available on Stellar.** There is not a required text book for this class. You can use Bodie, Kane, and Marcus as a reference. I will post the slides prior to each class. If you would like to take notes in the class, it might be a good idea to print a copy and bring it to the classroom. I will do my best to write a companion note such as this one for each class.

2 What do I Expect from the Students?

- **Come to the classroom with a love for the subject matter.** When I first realized that I could in fact make a career out of Finance, I was so exited that I put up a giant poster of the NYSE's trading floor in my tiny apartment. At the time, I was living in Brooklyn Heights and studying at NYU for my PhD in Physics. On my daily walk home via the Brooklyn Bridge, I often turned around at the end of the bridge to look at the lights in lower Manhattan, imagining myself to be in one of the office buildings. Looking back, it was all so laughable. The NYSE trading floor is definitely not where the action is taking place and I had never even been close to a Wall Street career. But that love for the subject matter is what really matters. Over the years, I have read many of the Finance books written for the popular press, like *Capital Ideas* by Peter Bernstein, *When Genius Failed* by Roger Lowenstein, *Liar's Poker* and *Flash Boys* by Michael Lewis, F̂ool's Gold by Gillian Tett. Many of the books, I simply cannot put them down after reading the first few pages. By contract, I have never read one popular Physics book even though I also have a PhD in Physics. I read Physics books only to prepare for exams. I often joke with my students that your real passion is reflected by the books on your night stand.

So if you read Finance books only for exams, maybe this is not your cup of tea. So why not quit Finance and look for your real passion? You are still young and have most of your life ahead of you. Why get yourself stuck in a career that you do not love? For those of you who have to take this class in the meanwhile, I will suggest that you fake it until you make it, or until the end of the semester, whichever comes first.

- **Be mentally present in the classroom.** Each morning when you get up, you might wash yourself, put on some clean clothes, and make yourself presentable. Have you ever thought about how to prepare your mind? When you eat, you might watch your diet and be careful with what you put into your mouth. Have you ever thought about what you feed to your mind? Your mind is the best instrument in your life, and yet, without proper training, your mind is also the most vulnerable, incapable of fending off the waves of distractions in this digital age. If you ever wonder why you are not doing well in the class, in an interview, or in your job, it is possible that you have not prepared your mind well. Being smart with a high IQ comes mostly from one's genetic inheritance. It is over-rated anyways. Instead, be impressed by people in your life who are mentally present and aware. They carry themselves differently and they are usually light-hearted, flexible, and happy.

  Being mindful is something you can develop. Why not start your practice in the classroom? Think of the classroom as a sanctuary, away from the digital distractions that permeate our society. Do you know that some people pay a lot of money to attend digital detox retreats just to get away from the Internet and Cell connections for a week? Here, you can do it for free. Just refrain from the Internet for 80 minutes. Trust me, the rest of the world will be just fine without your digital footprint. They probably won't even notice. Give yourself a break and give me your full attention during those 80 minutes. It will be a very good training of your mind.

- **As a general rule, all electronic devices including laptops and iPad are out.** If you really really need to use them, please talk to me. Please turn off your cell phones, including the *ding dong* sound for messages. If your phone is on vibration, please avoid putting it on hard surfaces. Otherwise, it will be just as noticeable.

- **Assignments and exams.** There are four group assignments, to be done in groups with no more than four students. Each assignment must be handed in before 5pm of the due day. Late assignments will not be accepted. The midterm exam will be given on Tuesday, October 18. The final exam will be given during the final exam week. There will be optional recitations held by TAs for the assignments and exams.

- **Let's keep our classroom a friendly environment.** This semester, we are going to spend 80 minutes together each time for 24 times. I would like us to create a friendly environment for one another. If one student happens to come into the classroom late, please try not to give him that "how dare you" look. Maybe he has a very good reason for being late. Who knows? Of course, this does not give you a free license to be late. You will certainly hear from me if I feel that you're consistently late. In this classroom, I would like everyone to feel comfortable enough to speak up, either for clarification or discussion. The only thing I would discourage in the classroom is students talking amongst themselves.

3 Modern Finance

After the 2008 financial crisis, I had the following conversation with my sister who works for a Pharmaceutical company. "Other than making money for themselves, what do people on Wall Street really do for the society?" She asked and I was quiet. "You know, everyday I go to work, I know that I am helping develop a drug that might help relieve people's pain, and I feel good about it. What about people on Wall Street?" She continued and I remained quiet. But inside, I had this huge question mark hanging over my head. To be honest, I have never fully resolved this doubt for myself. In 2008, the stock market dropped by 37%, and my passion for Finance was cut by just as much, if not more. What disappointed me was not the financial performance but the human performance in the financial industry.

In writing up the following account on the development of Modern Finance, I hope to be able to respond, at least partially, to my sister's question. Assuming that you will be part of the Wall Street in a few years, this should be a relevant question for you as well.

- **Markowitz (1952)** The beginning of Modern Finance can be dated by Markowitz (1952). As described in vivid details by Peter Bernstein in *"Capital Ideas,"* Harry Markowitz was a 25-year-old graduate student at Chicago, working on his PhD thesis. In 1990, he was awarded a Nobel Prize for his "pioneering work in the theory of financial economics."

  From the vantage point of today's knowledge base, the paper's insight is obvious and well understood by most people. First, Markowitz made the observation that, for any mean-variance investors, there is a risk and return tradeoff. Then, as any good researcher would do, he asked, given this tradeoff, what is this investor's optimal allocation to risk? "The answers Markowitz developed to these questions ultimately transformed the practice of investment management beyond recognition. They put some sense and some system into the haphazard manner in which most investors were assembling portfolios. Moreover, they formed the foundation for all subsequent theories on how financial markets work, how risk can be quantified, and even how corporations should finance themselves."[1]

  Against the backdrop of a single-minded focus on return at his time, Markowitz made the key insight that risk is central to the whole process of investing. In this day and age, any statement contrary to that observation is laughable. Risk is the single most important factor in Finance. No risk, no Finance. Financial markets, along with the tremendous innovations since 1970s, are vehicles designed to help us deal with risk.

  Well, this story does tell us how far Finance has evolved over the past half-century, with this one single insight made in the arcane world of academics by someone who had no direct interest or involvement in the stock market.

- **Tobin (1958)** Markowitz's insight was not recognized right away. After its initial publication in *the Journal of Finance*, the paper remained in obscurity for nearly ten years, attracting fewer than twenty citations in the academic literature. One of these citations was by James Tobin, a 1981 Nobel Prize winner. Tobin (1958) gave us the elegant result of two-fund separation. For any mean-variance investors, the optimal allocation consists of only two funds: one risky and one riskless. Regardless of their varying levels of risk aversion, all mean-variance investors hold exactly the same risky portfolio. The more risk-adverse investor allocates a smaller percentage of his wealth to the risky portfolio, but the composition of the his risky portfolio is exactly the same as everyone else.

[1] Chapter 2 of *"Capital Ideas,"* by Peter Bernstein.

This result gives us the striking insight that instead of getting lost in the sea of individual stocks, one should pay attention to this optimal risky portfolio. In today's world, with the increasing popularity of Index funds and ETFs, this idea seems quite obvious. But it was not until 1971, when the first Index Fund was created by John McQuown and his colleagues at Wells Fargo. And it was not until 1975, when the first Index Mutual Fund was created by John Bogle. This fund, now called the Vanguard 500 Index Fund, started off with just \$11.3 million, a 93% shortfall from the initial target of \$150 million. By 2014, over \$2 trillion is invested in index mutual funds, accounting for 20% of the total net assets of equity mutual funds. From 2007 through 2014, index domestic equity mutual funds and ETFs received \$1 trillion in net flows while actively managed domestic equity mutual funds experienced a net outflow of \$659 billion.[2]

Again, the influence from the academic world is unmistakable. If you read the stories surrounding the creations of these index funds, you will see that these pioneers in industry were often influenced at a personal level by a few professors. Their convictions were often strengthened by the intellectual power behind the academic research. John Bogle writes, "Nobel laureate economist Paul Samuelson played a major role in precipitating the index fund's creation... Samuelson was much more forceful, strengthening my backbone for the hard task that lay ahead: taking on the industry establishment. His article 'Challenge to Judgment' caught me at the perfect moment. Published in the inaugural edition of the Journal of Portfolio Management in the autumn of 1974, it pleaded that some large foundation set up an in-house portfolio that tracks the S&P 500 Index ... ."[3]

- **Sharpe (1964)** If you ask me to pick one model in Finance that has the biggest and the longest-lasting impact, it will be the CAPM. Following the stream of Markowitz (1952), Tobin (1958), and Sharpe (1964), one has the reaction that this sequence of intellectual development is so natural that it is inevitable. But the last step in this "Investment Trilogy" is truly a giant leap. In 1990, Bill Sharpe was awarded a Nobel Prize for his "pioneering work in the theory of financial economics."

  In Markowitz (1952), the attention is on an individual investor. How much he should include a stock in his portfolio is determined by this stock's contribution to the risk (variance) and return (mean) of his portfolio. As such, what matters are the correlations between this stock and all other existing stocks in the portfolio. In the CAPM, the attention is on the entire economy. If every investor behaves optimally according

[2] The 2015 Fact Book by Investment Company Institute (ICI). Posted on Stellar under Readings.
[3] "How the Index Fund Was Born," by John C. Bogle, *Wall Street Journal*, 2011.

to the calculation in Markowitz (1952), what happens to the entire market when you aggregate this optimal individual behavior? Collectively, the markets should also clear: borrowing and lending in the riskfree market must net out and the entire wealth of the economy must be 100% allocated to the risky portfolio. In the academic language, the CAPM is a result of taking the partial-equilibrium model of Markowitz (1952) to equilibrium, a beautiful insight from Economics.

In equilibrium, the optimal risky portfolio in Tobin (1958) becomes the market portfolio – the single most important factor in the economy. In such an economy, you no longer have to keep track of the correlations of one stock with respect to all other stocks. What matters is a stock's correlation with the market portfolio. Hence $\beta_i = \text{cov}(R_i, R_m)/\text{var}(R_m)$. In this way, CAPM further clarifies the concept of risk. In particular, risk is not measured by the variance of an individual stock. For two stocks with the same variance, the one that comoves a lot with the market portfolio is the more risky one. Why? Because risk that is not correlated with the market portfolio can be diversified, but there is no way to diversify away the risk in the market portfolio. This concept of systematic risk is by far the most important intellectual development in Finance.

After singling out the systematic risk, the equilibrium analysis gives us the elegant pricing result. Simply put, you get paid for bearing the risk that matters:

$$E(R_i) - r_f = \beta_i \left(E(R_m) - r_f \right) .$$

If a stock contains purely idiosyncratic risk with $\beta_i = 0$, then you do not get paid for bearing this risk and the expected return is the same as the riskfree rate: $E(R_i) = r_f$.

- **Black and Scholes (1973)** In the 1970s, Fischer Black, Myron Scholes and Bob Merton did their pioneering work on Continuous-Time Finance at MIT Sloan, on the second floor of E52, I was told. In 1997, Merton and Scholes were awarded a Nobel Prize for "for a new method to determine the value of derivatives."

  The impact of this work is such that taking it out would be like switching off a bright light and the world of Finance would be a dim field due to its absence. Many of the financial markets we are going to study in this course would not have been created without this work. Even if such markets were in existence, people in these markets would be having a hard time figuring out how to price the product or hedge the risk.

  Going back to my sister's question, "What does Finance do for the society?" Finance helps people deal with risk. No risk, no Finance. Markowitz (1952) gives us a framework to quantify the risk and return tradeoff. Sharpe (1964) points out that not all risk is equal and only systematic risk should be compensated. The work by Black, Merton, and Scholes takes this business of risk to a whole new dimension. When you purchase a stock, you get the entire package of risk that is inherent in this stock. In the language of academic Finance, you have a linear position and you own the entire distribution of the risk, both up and down. What if you don't want the entire distribution? What if you are interested in taking only some of the risk but not all? The financial innovation inspired by the work of Black, Merton, and Scholes is all about giving you more flexibility in dealing with risk.

Every summer, I go back to Shanghai to spend a month and half with my parents. In May 2015, before I was able to purchase the air ticket, I had to wait for a test result from my doctor. But it was getting close to the departure date and I was anxious that the airfare might jump up. So what did United Air offer me in this situation? A fare lock. I paid \$7.99 to have a 7-day option to purchase the ticket at the prevailing price on that day regardless of how the price might fluctuate over the following week. For a 72-hour lock, they charged \$5.99. As I was first writing this lecture note in August 2015, just out of curiosity, I checked the price again. The 7-day lock cost \$11.99. So I inferred that airfare must have turned more volatile from May to August 2015. As a matter of fact, I was tempted to write a code to automatically collect the fare lock price once a day so as to back out the pricing model. Is the society better because of this product? At least I was able to wait for my test result without the added anxiety about airfare fluctuations.

Sure, this is not a financial product but the underlying message is the same. In the presence of risk, it helps if you could give people more flexibility in the kind of risk they take. This example is simpler and easier to communicate. But financial products serve the same purpose for individuals and corporations. So why are the general public so negative about financial innovation? The former Fed Chairman Paul Volcker was quoted in saying that the most important financial innovation that he has seen in the past 20 years is the automatic teller machine. He then added that this is more of a mechanical innovation than a financial one. The practices of some financial institutions and individuals deserve 100% of the criticism. There is not question about it. But, in my personal opinion, the criticism piled up on the innovation itself is perhaps misplaced.

4 Financial Markets

I grew up in Shanghai with very little knowledge about stocks or bonds. It's possible that these names had never even showed up in my vocabulary. At that time in China, people read Philosophy books and looked up to scientists. My dream was to become Marie Curie. Dealing with money was just beneath me. In November 1990, several months after I left Shanghai for the US, the Shanghai Stock Exchange was re-established. The rest, as they say, is history. It is probably not an exaggeration to say that a one-day coverage of China in the *Wall Street Journal* today equals that of a year back in 1990. In any case, when I finally concluded that Physics and myself had no future together, I was already in New York. A friend told me about Finance and recommended me to read Bernstein's book. For someone with absolutely no knowledge about Finance, it was truly an eye-opener. Let me borrow his words in describing the financial markets:

> Financial markets are among the most dazzling creations of the modern world. Popular histories of financial markets from the City of London to Wall Street tell the story of panics, robber barons, crooks, and rags-to-riches tycoons. But such colorful tales give little hint of the seriousness of the business that goes on in those markets. John Maynard Keynes once remarked that the stock market is little more than a beauty contest and a curse to capitalism. And yet no nation that has abandoned socialism for capitalism considers the job complete until it has a functioning financial market.
>
> Simply put, Wall Street shapes Main Street. It transforms factories, department stores, banking assets, film producers, machinery, soft-drink bottlers, and power lines into something that can be easily convertible into money and into vehicles for diversifying risks. It converts such entities into assets that you can trade with anonymous buyers or sellers. It makes hard assets liquid, and it puts a price on those assets that promises that they will be put to their most productive uses.
>
> Wall Street also changes the character of the assets themselves. It has never been a place where people merely exchange money for stocks, bonds, and mortgages. Wall Street is a focal point where individuals, businesses, and even entire economies anticipate the future. The daily movements of security prices reveal how confident people are in their expectations, what time horizons they envisage, and what hopes and fears they are communicating to one another.[4]

[4] Introduction of "Capital Ideas," by Peter Bernstein.

Over the next few months, I hope to be able to teach financial markets with this sense of awe, which I felt when I first learned about them. At a personal level, I am not too involved with the markets. Whatever money I have, I put them in index funds and wish to see them again after retirement in a few years. Buy-and-forget pretty much sums up my investment strategy. And yet, I follow the development of the markets with great interest. I view it as a stage with great drama. There are uncertainty, human emotion, and, surprisingly, logic, rationality and regularity. In this course, we will cover the three key markets: equity, fixed-income, and derivatives. More specifically, we will study in depth the US equity markets, the equity index options, US Treasury bonds, corporate bonds, interest-rate swaps, and credit derivatives. If time permits, we will also cover currency. In the next section, I'll explain the topics to be covered in more detail.

5 Topics to be Covered

As an academic discipline, what Finance can offer to financial markets is a "cool head." A roller-coaster ride might be exciting initially, but after a few ups and downs, anyone with a sensible mind would ask, is there more to it? Likewise, when being buried deeply in the ups and downs of financial markets, most of us welcome the opportunity to extricate ourselves from the noise and busyness to get a better view. We are often told to learn from our own experiences in life. Empirical Asset Pricing is about learning from the experiences in the financial markets. For this, we have a list of quantitative tools and models at our disposal, which have been developed and widely used. By now, these tools and models have become part of the language on Wall Street. Warren Buffett might have the luxury of not knowing them or even making fun of them. But before you become him, you probably need to know. In any case, these are fun tools and models to learn, especially when you apply them to real financial data.

Let me list the topics to be covered over this semester. I've put a class number to each topic so as to have some discipline. But once in a while I might have to modify our schedule. By the end of the semester, if we somehow end up with extra time, I hope to be able to cover market micro-structure with topics like price discovery, information trading, market making, and high-frequency trading.

1. **Introduction: Class 1**.

2. **Equity:**

   (a) **Class 2**. Alpha and Beta.

(b) **Classes 3 & 4**. Equity in the Cross-Section, Fama-French Three-Factor Model.

(c) **Classes 5 & 6**. Other Cross-Sectional Trading Strategies and Currency Carry Trades.

(d) **Class 7**. Equity in the Time-Series, Time-Varying Expected Returns.

(e) **Classes 8 & 9**. Equity in the Time-Series, Time-Varying Volatility.

3. **Option:**

(a) **Class 10**. Option: Introduction and the Black-Scholes Model.

(b) **Class 11**. **MidTerm Exam.** (covers Classes 1 to 9.)

(c) **Class 12 & 13**. Option: Model to Data, Volatility Smirks and Tail Risk.

(d) **Class 14**. Option: Beyond the Black-Scholes Model.

4. **Special Topic:**

(a) **Classes 15 & 16**. Risk Management.

5. **Fixed Income:**

(a) **Class 17**. Bond: Yield and Duration.

(b) **Class 18**. Bond: Yield Curve.

(c) **Class 19**. Bond: Term Structure Models.

(d) **Class 20**. Bond: Interest Rate Swaps.

(e) **Class 21**. Credit: Corporate Bonds and the Merton Model.

(f) **Class 22**. Credit: Credit Default Swaps and Other Models of Default.

6. **Portfolio Management:**

(a) **Class 23**. The Process of Portfolio Management and Optimal Risky Portfolio.

(b) **Class 24**. The Black-Litterman Asset Allocation Model.

6 Quantifying Risk

A professor went to a Zen master, inquiring about Zen. The master invited him to sit down and have a cup of tea. When the tea kettle arrived, the master started to pour hot water into the cup. Soon, the cup was full, but the master kept on pouring. The professor exclaimed, "Master, master, the cup is full. It cannot take any more water!" The master responded, "That's right, Professor. Like this cup, your mind is full of your own views and opinions. How can I teach you Zen?"

So whatever opinion and knowledge you might have about what I am going to teach, please throw them out and keep your mind open. You might be able to do a histogram with eyes closed. You might know the pdf and cdf of a normal distribution inside out. Still, your understanding might be lacking. Going through the motions is easy. Understanding the insight takes more attention, patience, and reflection. Just like anything in life, it is the insight that really matters. Knowledge without insight is dead knowledge.

6.1 Data

To understand what uncertainty means, let's start with the past experiences of uncertainty in the US stock market. The time-series plotted in Figure 1 contains the annual stock returns in the US markets from 1927 through 2015. It uses the CRSP value-weighted index, which includes all stocks traded on the three major US exchanges (NYSE, AMEX, and NASDAQ). It is an index preferred by academics. The reported returns are calculated from year-end to year-end, including both capital gains and distributions (i.e., dividends). Another index that would work equally well to represent the overall market is the S&P 500 index. In other words, from a value-weighted perspective, the entire stock market can be very well captured by those 500 large-cap stocks included in the S&P 500 index.

Later in this class, we will sample the data at a daily frequency, and things will look very different. I am using the annual sampling frequency as an example here because it is a widely used horizon. (Also, it keeps our example simple with relatively small amount of data points.) In assessing the performance of a stock or a managed portfolio, the average annual return is often the first benchmark. As such, you should be very comfortable with comparing performance numbers at this frequency.

As a first step, study the plot, follow the ups and downs, absorb the information at an intuitive level. Be curious, and ask yourself questions. This is your random walk down wall street. For example, as an exam question, I could ask you, in the history of the US stock market, what were the worst one-year returns? When did they happen? I might not expect

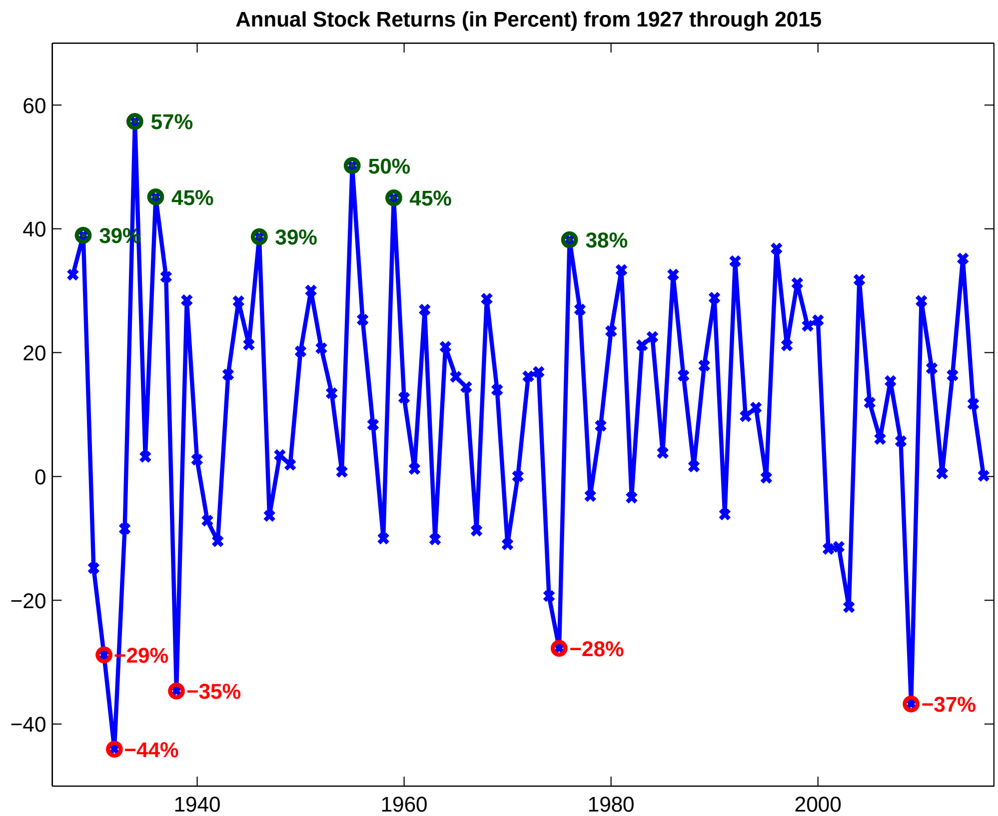

Figure 1: Time-series of annual stock market returns from 1927 through 2015. Returns are calculated using the CRSP-value weighted index, which includes all stocks traded on the US exchanges. Source: Prof. Ken French's website.

you to be so accurate as to tell me -44% in 1931 or -37% in 2008. But I would expect you to know that it was closer to -40% than -20% or -60%. I also would expect you to know about the great depression in the 1930s and the more recent crisis in 2008. Moreover, if I ask you, do you consider a 20% annual return to be normal in the US stock market? You should be ready to articulate your response with the help of the time-series data. Or suppose I tell you that the Shanghai Stock Exchange (SSE) composite index was at 2,115.98 at the end of 2013 and increased to 3,234.68 at end end of 2014. Using the US experiences, how unusual is such an annual performance? These are not hard questions once you really get yourself into the data.

Don't force yourself to memorize these numbers. You will soon forget. What is the point? You need to be interested. You don't need to force yourself to memorize your hometown, your good friends, or your family, do you? I once read someone describing how it was like working with Robert Rubin, the former Treasury Secretary and a long-time executive in Goldman Sachs. I don't remember the exact description but the impression I had was that, when presented with data and plots, this guy would just dive into them with his pencil and mind. Suppose you are presented with a plot like Figure 1 for the first time. And you react with folded arms and the look of "what is the big deal" and "please impress me," then you probably are not going to enjoy Finance as a profession.

Finally, if we look more closely, we will also notice that the large movements are related to the economic conditions. Using the business cycles dated by the NBER, we can see that the largest five negative returns all happened during recessions. Both 1931 (-44%) and 1930 (-29%) fell in the middle of the severe recession from August 1929 to March 1933; 2008 (-37%) was in the recession from December 2007 to June 2009; 1937 (-35%) in the recession from May 1937 to June 1938; and 1974 (-28%) in the recession from November 1973 to March 1975.[5] On paper, these events might be distant and without any real connection. But if you had to live through any of these crises, the impact would be totally different. For example, Warren Buffett was born in August 1930, ten months after the great crash of 1929 and right in the midst of the great depression. His dad, being a stock broker for the Union State Bank, was having trouble feeding the family.[6] Clearly, an upbringing like this helps shape one's investment philosophy and, more broadly, one's attitude toward life. A trader/investor who has lived through a crisis like 2008 would look upon the financial markets with a new set of eyes. So if you would like to become a great investor, try to make a closer connection to the historical events, through books, newspapers, and accounts by older and wiser people. Make these events real for yourself and they would be a treasure sitting in the background,

[5]For a complete list of the NBER business cycles, please go to `http://www.nber.org/cycles.html`.

[6]See *"The Snowball: Warren Buffett and the Business of Life,"* by Alice Schroeder.

emerging when the moment calls for them. After all, life is an endless cycle of ups and downs. So are the financial markets.

6.2 Histogram

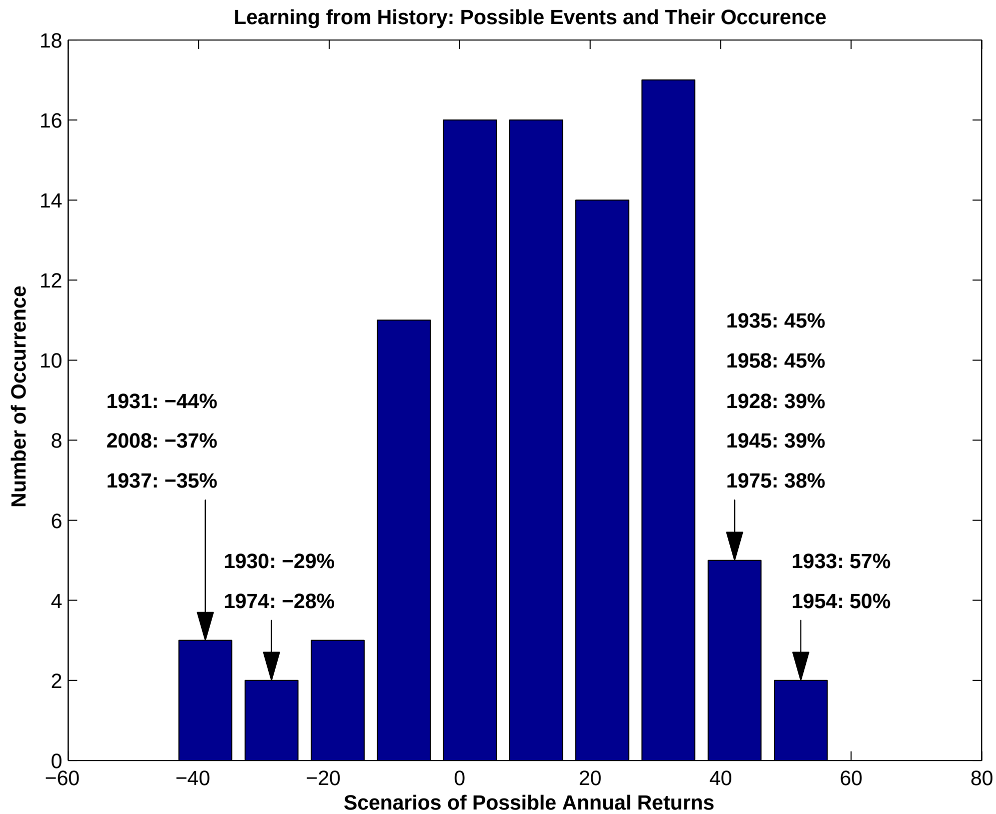

Figure 2: Histogram of stock market returns from 1927 through 2015.

A histogram is like a CT scan on financial returns. In a CT scan, all parts of an image are captured at the same time. In a histogram, however, the image is a compilation of data points collected over a long time span — in some cases, many decades. This is because the financial markets, especially the equity markets, are known to be volatile. So with a few years of data, all you see is noise; but with years of repetition, the noise gets washes out and the valuable signal emerges with more precision. So doing a histogram is like pointing a camera at the stock market for a really really long time. In our current example, the camera has been on for 89 years from 1927 to 2015.

Comparing Figures 1 and 2, you notice that a histogram transforms the dynamic time-series data into a static one. In creating the histogram in Figure 2, we do not care at all about the sequence of the events – which data point happens first and which one happens later. All we need are the outcomes: the 89 data points, i.e., 89 returns. We line up the returns from the worst (-44.04%) to the best (57.35%), chop the entire interval evenly into N bins (N=10 in our example), with each bin the size of (0.5735+0.4404)/10=0.1014. We then count how many data points (i.e., returns) fall into each bin. So that's a histogram. Sounds really easy, and it is indeed very easy. What is the big deal?

Conceptually, the transformation from Figure 1 to Figure 2 is a significant one. It changes how we view the data. In Figure 1, you are bombarded with uncertainty. In Figure 2, we step out of it and begin to deal with it, by developing a *regularity* about the uncertainty. In order to do so, however, we have to make some non-trivial assumptions. In creating the histogram in Figure 2, the underlying assumption is that every year from 1927 to 2015, the uncertainty in the stock market is exactly the same. In other words, sitting in the background is a distribution machine (like the wizard of oz), which spits out returns with a regularity (i.e., likelihood) over a certain range. Every year, this distribution gives you one realization. You record the outcome. One year later, it totally forgets what it gave you the year before and draws from that identical distribution one more time and give you a new realization. You dutifully record the outcome. After 89 years of repetition, you tell yourself, I should be smart about it. After all, it is going to draw from that same distribution machine again. Let me use the 89 data points I have recorded so far to paint a picture of that distribution. Out comes the histogram in Figure 2. Now, instead of swimming in the sea of uncertainty like in Figure 1, you are looking at the uncertainty from the lens of Figure 2. The future is still uncertain, but you are armed with a tool to deal with the future uncertainty. You know what to expect.

Of course, this is only true if the world functions in this forgetful, and yet consistent way. In Statistics, this underlying assumption is called *iid*: stock returns are independent (forgetful) and identically (consistent) distributed. It is exactly because of this *iid* assumption, we are willing to throw away the sequencing information in doing the histogram and focus only on the outcomes. Otherwise, the analysis will be done differently. Suppose that the real world is not *iid*. Instead, one year of good performance is more likely to be followed by another year of good performance. Then for sure, we will not be throwing away the sequencing information. Later in the semester, we will discuss this possibility of predictive returns. Throughout the semester, you will notice that in Finance we often make strong assumptions first, and examine the markets under these assumptions. Then we realize that perhaps the initial assumptions are too strong. We then relax the assumptions and re-examine the markets. This is the typical process of an empirical investigation. Strong assumptions are never the problem, but making financial decisions under some assumptions and yet not being aware of them is often the problem.

At this point, let me summarize the things we know and the things we are not sure about. Let's suppose that now is year t and let's use $\tilde{R}_{t+1}$ to denote the stock returns over the next year. Notice that $\tilde{R}_{t+1}$ a random variable. I put a tilde on top of it to remind you that it is not a number. Associated with this random variable is the histogram in Figure 2, which tells you all of the possible scenarios and their likelihood. This is what we know about about $\tilde{R}_{t+1}$: the future outcome will be drawn from a distribution centered around a value, which we call the expected stock return $E(\tilde{R}_{t+1})$. Taking the average of the 89 data points, we say that $E(\tilde{R}_{t+1})$ can be approximated by the historical average of 12%. The eventual outcome of $\tilde{R}_{t+1}$ remains uncertain. A year later, we will get to see its realization. It could be something like 2008 (-37%), or it could be something like 1954 (50%). *Ex ante* (i.e., before the fact), we can tell you the probability of such outcomes. But *ex post*, what will eventually happen? We are not sure. This is the uncertainty faced by everyone, no matter how powerful that person might be. Maybe it is to express the frustration over the lack of knowing, we call this uncertainty *risk*.

The limitation of such an empirical exercise of learning from the history is the history itself. If a certain kind of risk has not yet happened, then it will not be part of our histogram. For example, before the S&P 500 index dropped by -20% over just one day on October 19,1987, this kind of one-day event was not in anyone's histogram. Now that it happened, it adds to the left tail of the daily distribution and our "model" is updated.

6.3 Probability Density Function (PDF)

The histogram in Figure 2 tells us how often an event falls within a certain bin. It is not a probability distribution yet because a probability distribution needs to add up to one. From Figure 2 to Figure 3, the shape of the distribution remains exactly the same, but the labeling of the y-axis has changed from "number of occurrences" to "probability density." If you go over the Matlab code attached in the end of this note, you will see that I scaled the entire plot by a constant: 0.1014×89, where 0.1014 is the width of the bin and 89 is the total number of events. This way, the entire area of the blue bars sums up to one. The probability of all likely events adds up to one. A histogram now becomes a probability density function. We call it an empirical pdf because it builds from the data. Now let's introduce a model to mimic this empirical distribution.

Carrying around an empirical distribution is cumbersome. The next step is to introduce a well behaved analytic distribution to approximate it. This is where normal distribution

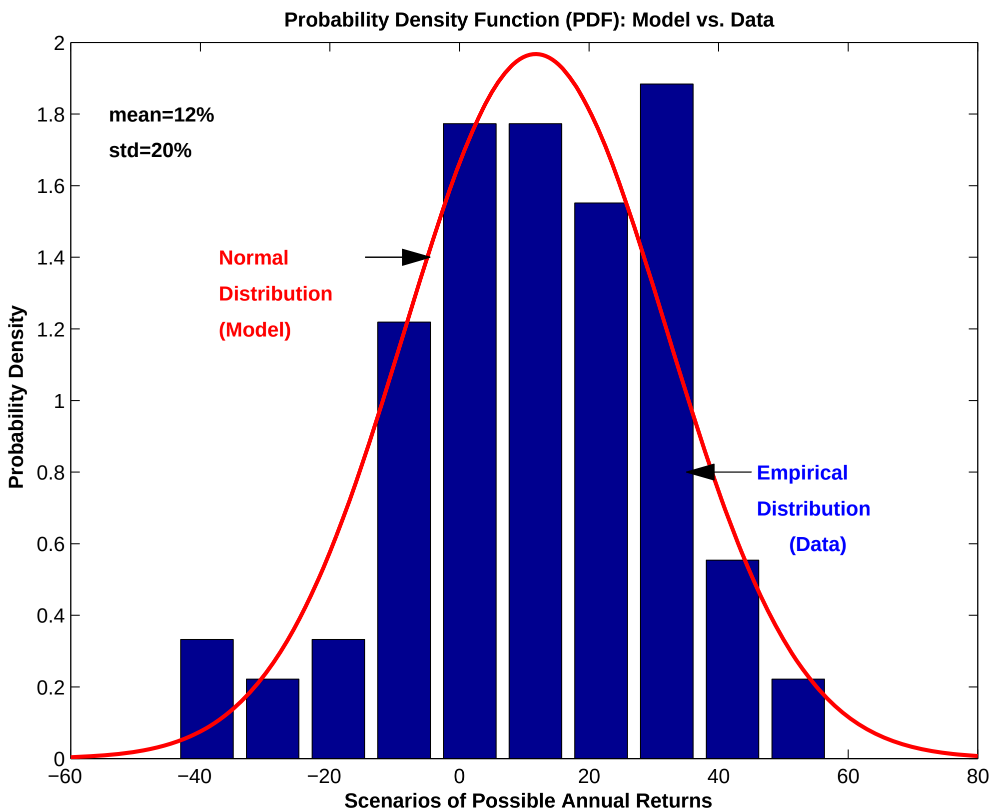

Figure 3: Probability density function, model vs. data.

comes in. You will see that normal distribution plays a very important role in Finance. Let's start with the probability density function of a normal distribution with mean μ and volatility σ:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

If this looks discouraging, don't be. Because you can do `normpdf(x,`μ`,`σ`)` in Matlab and out comes the value of $f(x)$. In Excel, there is something that is just as easy. Somehow, computers make math more accessible. Those people who are good at equations suddenly become less attractive. In any case, this is how I plotted the red line in Figure 3. I first calculated the mean (approximately 12%) and standard deviation (approximately 20%) of my 89 data points. Then, for each value in the of stock returns, say -44%, I plugged in `normpdf(-0.44,0.12,0.20)` to get its pdf value, without having to remind myself the exact expression of a Gaussian function.[7]

Let's now formalize what we have done so far. Out of the historical experiences of 89 returns, we are able to develop the following regularity about the risk in the stock market. Let $\tilde{R}_{t+1}$ be the uncertain return over the next year. Learning from the history, we know that its expected value is $\mu = 12\%$ and its standard deviation (i.e., volatility) is $\sigma = 20\%$. Moreover, we know that its distribution can be approximated by $f(x)$. All of this can be summarized by the following model of stock returns:

$$\tilde{R}_{t+1} = \mu + \sigma \, \tilde{\epsilon}_{t+1}$$

where $\tilde{\epsilon}_{t+1}$ is a standard normal distribution: $E(\epsilon) = 0$ and $\mathrm{std}(\epsilon) = 1$.

We do not have time to formally test the model, but from Figure 3, it seems to be a reasonable approximation of the data. Not perfect, but reasonable and useful, as you will see. But as useful as a model might be, you should always use it with extreme caution. The limitation of a model is two fold. First, it builds from the empirical distribution, which itself is limited by our experiences. Second, it is an analytic approximation of the much richer reality. Some of the financial crises happened exactly when investors become too comfortable with their models in their pricing, hedging, and trading.

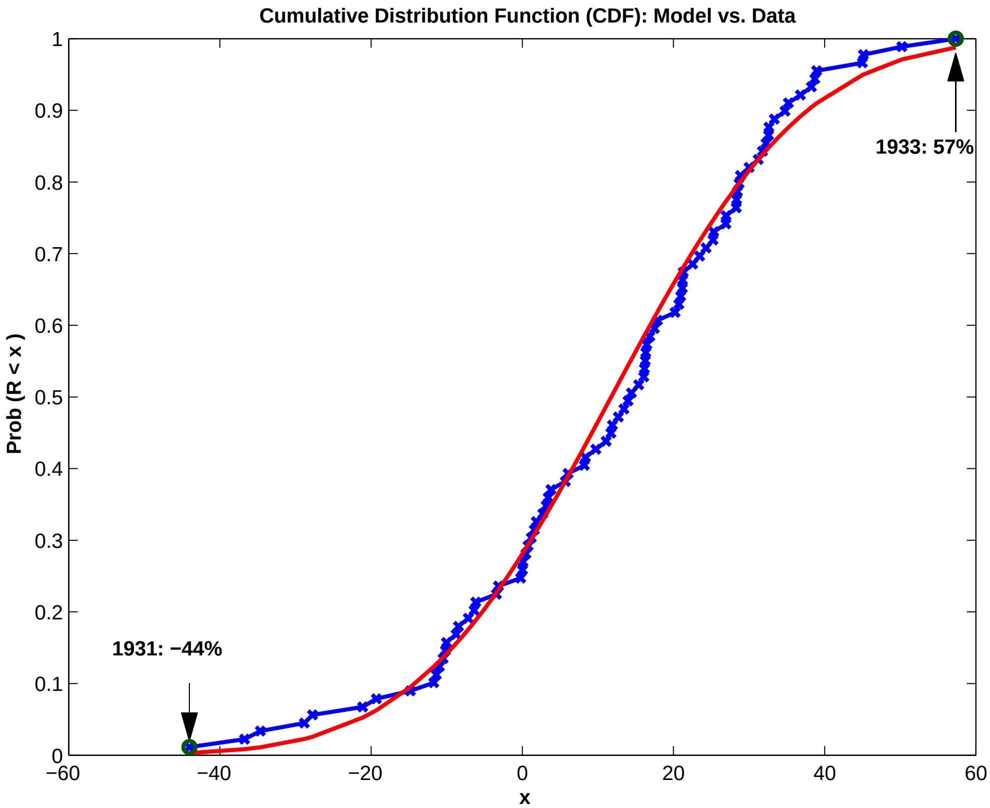

Figure 4: Cumulative distribution function, model vs. data.

6.4 Cumulative Distribution Function (CDF)

In order to plot the blue line in Figure 4, I line up the 89 data points from left to right, from the most negative to the most positive. So -44% in 1931 is my leftmost data point, with -37% in 2008 standing next to it, and 57% in 1933 is my rightmost data point. There are 89 points in total. So the values of the empirical CDF are:

$$\begin{aligned}
&\text{Prob}(\tilde{R}_{t+1} \leq -0.44) &&= 1/89 \\
&\text{Prob}(\tilde{R}_{t+1} \leq -0.37) &&= 2/89 \\
&\ldots \\
&\text{Prob}(\tilde{R}_{t+1} \leq 0.57) &&= 89/89
\end{aligned}$$

That is how the blue line is constructed. Isn't it simple? Now we have an empirical CDF.

[7]To be more precise, I use `normpdf(-0.44,mean(RM/100),std(RM/100)`, where RM is the time-series of the 89 market returns in percentage and mean and std are the Matlab functions to calculate mean and standard deviation. See the Matlab code in the Appendix.

The CDF for the model is just as simple. Mathematically, for any value x, it is

$$\text{Prob}\left(\tilde{R}_{t+1} \leq x\right) = \int_{-\infty}^{x} f(z)\, dz\,,$$

which might look intimidating until you realize that you can use Matlab and do:

$$\text{Prob}\left(\tilde{R}_{t+1} \leq x\right) = \texttt{normcdf}(\texttt{x}, \mu, \sigma)\,.$$

This is how I plotted the red line, I plugged all 89 returns into `normcdf` with $\mu = 12\%$ and $\sigma = 20\%$.[8] That is why the red line is not as smooth as a textbook example: the 89 data points are not evenly spaced. For example, to move from -44% to -37% is a pretty large gap. In my program, I do normcdf on these two returns and ask Matlab to link the two points with a straight line. To get a smoother line, I could have asked the normcdf to evaluate many more returns in between -44% and -37% and then connect the points. This is the advantage of a model over data. The data only gives us two experiences, -44% and -37%, and nothing in between. But the model can extrapolate to places that experiences did not take us. Needless to day, this is always the danger of a model, especially when we mistake the model as the reality.

6.5 Models are Limited

The normal distribution model is a corner stone of Finance. In the CAPM, we care only about the first two moments (mean and variance) of stock returns. Implicitly, we are treating returns as a normal distribution with μ and σ. In the Black-Scholes model, this assumption of normal distribution is explicitly given. That is why in the pricing formula, you see normcdf here and there. This model offers simplicity and elegance, a great vehicle to carry us far. But once in a while, it carries us too far.

Let's take a look at an example when this model fails. Figure 5 plots the daily returns of the S&P 500 index. Comparing with the time-series in Figure 1, we have more data points and the returns fluctuate within a much narrower band. Over an year, the volatility is about 20%. Over a day, it is about 1%.

Using the daily data, Figure 6 plots the CDF for the left and right tails. The red line is the CDF produced by normal distribution, with the same mean and variance as the empirical distribution, and the blue and green lines are the CDFs produced by the historical

[8] Again, for illustration purpose, I am using 12% and 20% as the sample mean and standard deviation. In practice, the actual sample mean and standard deviation for the 89 data points are used.

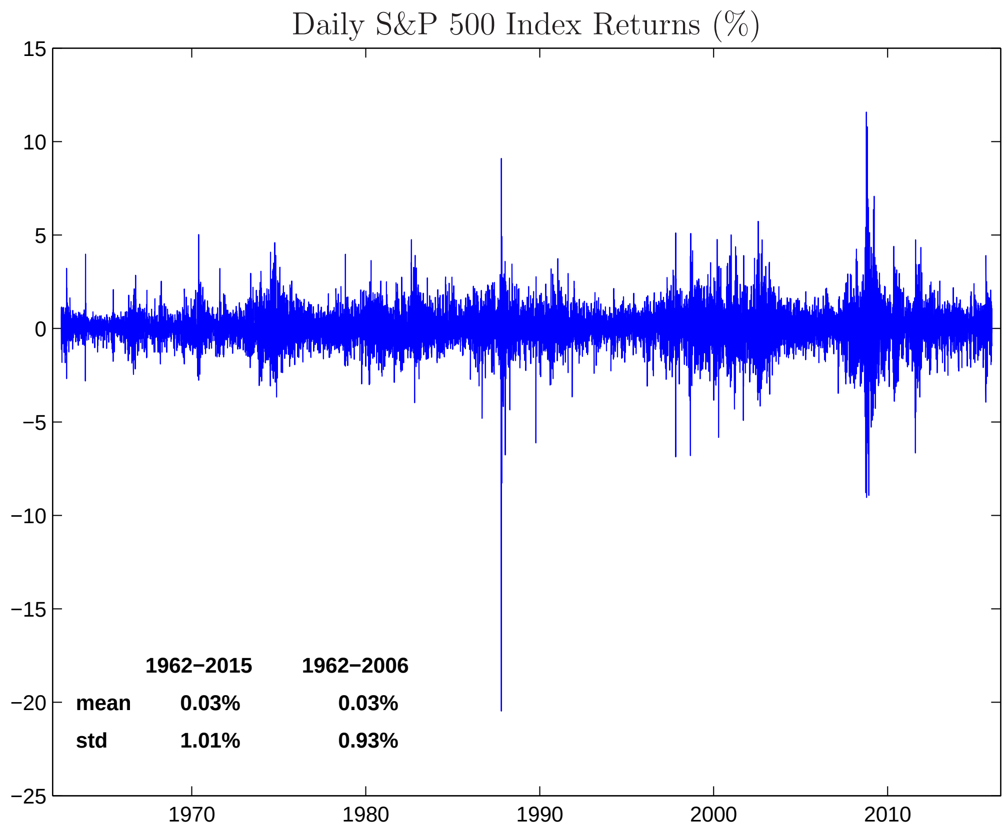

Figure 5: Time-series of daily returns on the S&P 500 index from 1962 through 2010. Source: Wharton Research Data Service (WRDS).

experiences. As it is evident in the plots, both the left and right tails are much thicker in the data. If your financial instruments are very sensitive to the tails (e.g., a far out-of-the-money put option), then this difference would really matter. In fact, this is one of the first places where the Black-Scholes model fails to perform.

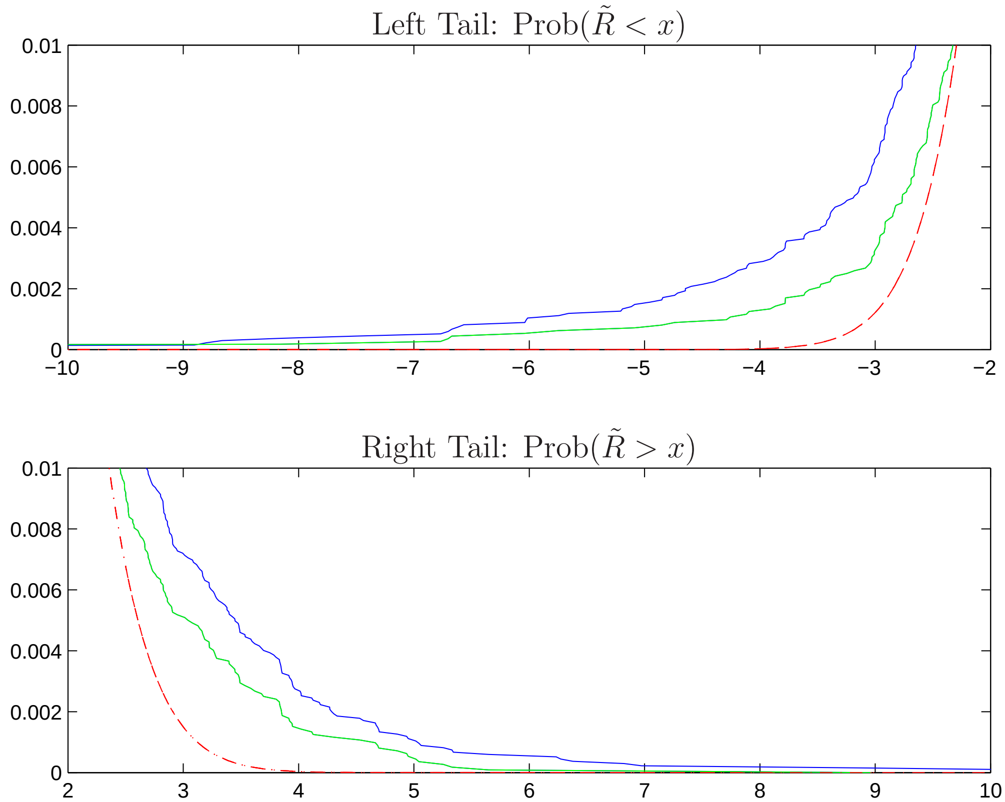

Figure 6: Cumulative distribution function using daily returns, model vs. data. The x-axis measures the daily standard deviation moves: daily returns normalized by the sample standard deviation. The y-axis measures the cumulative probability: 0.01 is 1%. The blue line is the empirical CDF generated by daily returns from 1962 through 2015. The green line is the empirical CDF generated by daily returns from 1962 through 2006. The red line is the analytic CDF generated by a normal distribution with the same mean and standard deviation as the entire sample (1962-2015).

A Matlab Code

Code 1: Histogram.m

```
load FF_Factors_Annual.txt; % Posted on Stellar under Data. Orignally
    from Prof. Ken French's website.
Data=FF_Factors_Annual; %Data Content: year, Mkt-RF, SMB, HML, RF

YR=Data(:,1); RF=Data(:,end); RM=Data(:,2)+RF; %RM=Mkt, RF=riskfree
Time=datenum([YR 12*ones(length(YR),1) 31*ones(length(YR),1)]); %I always
    prefer to work with the Matlab time. All returns are realized
    at the year end.

% Plot the time-series data
figure(1); clf;
h=plot(Time,RM,'bx-','LineWidth',2);
datetick('x','yyyy');
title('\bf Annual Stock Returns (in Percent) from 1927 through 2015');
hold on;

% List the worst five years
[tmp,Index]=sort(RM); Worst=[]; Best=[];
for i=1:5,
  Worst=[Worst; [year(Time(Index(i))),round(RM(Index(i)))]];
  p=text(Time(Index(i))*1.0005,RM(Index(i)),['\bf' num2str(round(RM(Index(i)
      ))) '%']);
  set(p,'Color','red');
  plot(Time(Index(i)),RM(Index(i)),'ro','LineWidth',2);
end
% List the best seven years
for i=0:6,
  Best=[Best; [year(Time(Index(end-i))),round(RM(Index(end-i)))]];
  p=text(Time(Index(end-i))*1.0005,RM(Index(end-i)),['\bf ' num2str(round(RM
      (Index(end-i)))) '%']);
  darkgreen=[0 0.35 0];
  set(p,'Color',darkgreen);
  plot(Time(Index(end-i)),RM(Index(end-i)),'Marker','o','Color',darkgreen,'
      LineWidth',2);
end
```

```
hold off;
axis([datenum([1925 12 31]) datenum([2016 12 31]) -50 70]);

% Histogram
figure(2)
[Occurence Scenario]=hist(RM);
bar(Scenario,Occurence);
axis([-60 80 0 18])
xlabel('\bf Scenarios of Possible Annual Returns');
ylabel('\bf Number of Occurrence');
title('\bf Learning from History: Possible Events and Their Occurence');
hold on;

% Mark the left tail
text(-55,9,['\bf ' num2str(Worst(1,1)) ': ' num2str(Worst(1,2)) '%']);
text(-55,8,['\bf ' num2str(Worst(2,1)) ': ' num2str(Worst(2,2)) '%']);
text(-55,7,['\bf ' num2str(Worst(3,1)) ': ' num2str(Worst(3,2)) '%']);;
arrow([Scenario(1) 6.5],[Scenario(1) 3]);
text(-37,5,['\bf ' num2str(Worst(4,1)) ': ' num2str(Worst(4,2)) '%']);
text(-37,4,['\bf ' num2str(Worst(5,1)) ': ' num2str(Worst(5,2)) '%']);
arrow([Scenario(2) 3.5],[Scenario(2) 2]);

% Mark the right tail
text(50,5,['\bf ' num2str(Best(1,1)) ': ' num2str(Best(1,2)) '%']);
text(50,4,['\bf ' num2str(Best(2,1)) ': ' num2str(Best(2,2)) '%']);
arrow([Scenario(end) 3.5],[Scenario(end) 2]);
text(40,11,['\bf ' num2str(Best(3,1)) ': ' num2str(Best(3,2)) '%']);
text(40,10,['\bf ' num2str(Best(4,1)) ': ' num2str(Best(4,2)) '%']);
text(40,9,['\bf ' num2str(Best(5,1)) ': ' num2str(Best(5,2)) '%']);
text(40,8,['\bf ' num2str(Best(6,1)) ': ' num2str(Best(6,2)) '%']);
text(40,7,['\bf ' num2str(Best(7,1)) ': ' num2str(Best(7,2)) '%']);
arrow([Scenario(end-1) 6.5],[Scenario(end-1) 5]);
hold off

% Plot the distribution, pdf
figure(3); clf;
D_X=mean(diff(Scenario/100)); %Scenario and RM are in return space. To
    be careful, I always do the math in decimal.
```

```
N_norm=D_X*sum(Occurence);
bar(Scenario,Occurence/N_norm); %rescale the number of occurence to get
     probability density.
hold on;
PDF=normpdf(-0.60:0.01:0.80,mean(RM/100),std(RM/100)); %probability
    density function.
h=plot(-60:1:80,PDF,'r-','LineWidth',2);
hold off
V=axis; axis([-60 80 V(3:4)]);
text(-38,1.4,'\bf Normal','Color','r');
text(-38,1.3,'\bf Distribution','Color','r');
text(-38,1.2,'\bf (Model)','Color','r')
arrow([-14.5 1.4],[-4.5 1.4]);
text(45,0.8,'\bf Empirical','Color','b');
text(45,0.7,'\bf Distribution','Color','b');
text(50,0.6,'\bf (Data)','Color','b')
arrow([45 0.8],[35 0.8]);
text(-55,1.8,['\bf mean=' num2str(round(mean(RM))) '%']);
text(-55,1.7,['\bf std=' num2str(round(std(RM))) '%']);
xlabel('\bf Scenarios of Possible Annual Returns');
ylabel('\bf Probability Density');
title('\bf Probability Density Function (PDF): Model vs. Data');

% Plot the distribution, cdf
figure(4); clf;
mean_R=mean(RM/100); std_R=std(RM/100); %to be safe, always do my math
    in decimal.
sorted_RM=sort(RM); Y_vec=cumsum(ones(length(RM),1))/length(RM);
Y_norm=normcdf((sorted_RM/100-mean_R)/std_R); %again, do my math in
    decimal.
plot(sorted_RM,Y_vec,'b-x',sorted_RM,Y_norm,'r-','LineWidth',2);
hold on
title('\bf Cumulative Distribution Function (CDF): Model vs. Data');
ylabel('\bf Prob (R < x )');
xlabel('\bf x');
% Mark the left tail
text(-55,0.15,['\bf ' num2str(Worst(1,1)) ': ' num2str(Worst(1,2)) '%']);
arrow([sorted_RM(1) 0.1],[sorted_RM(1) 0.02]);
```

```
plot(sorted_RM(1),Y_vec(1),'o','Color',darkgreen,'LineWidth',2);

% Mark the right tail
text(46,0.85,['\bf ' num2str(Best(1,1)) ': ' num2str(Best(1,2)) '%']);
arrow([sorted_RM(end) 0.87],[sorted_RM(end) 0.98]);
plot(sorted_RM(end),Y_vec(end),'o','Color',darkgreen,'LineWidth',2);
hold off;
```