

15.433 INVESTMENTS

Class 2:

Securities, Random Walk on Wall Street

Reto R. Gallati

MIT Sloan School of Management

Spring 2003

February 5th 2003



Outline

Probability Theory

- A brief review of probability distributions
- Evaluating random events with normals.
- Large surprises and normal distributions.

Statistical Data Analysis

- Empirical distributions. Sample statistics.
- The precision of sample statistics.

Summary

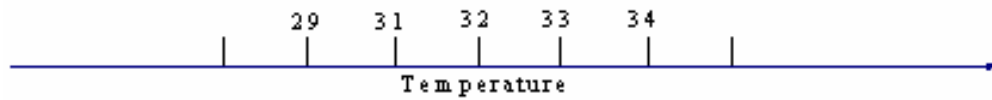
Questions for Next Class

What makes an event random

- Flipping a Coin:



- Forecasting tomorrow's temperature



Probability Distributions

Mathematical tools for random events. It has two components: outcomes and their likelihood.

Examples:

- Binomial Distribution

$$\tilde{X} = \begin{cases} 0 & \text{with probability } p \\ 1 & \text{with probability } 1-p \end{cases}$$

- Standard Normal Distribution

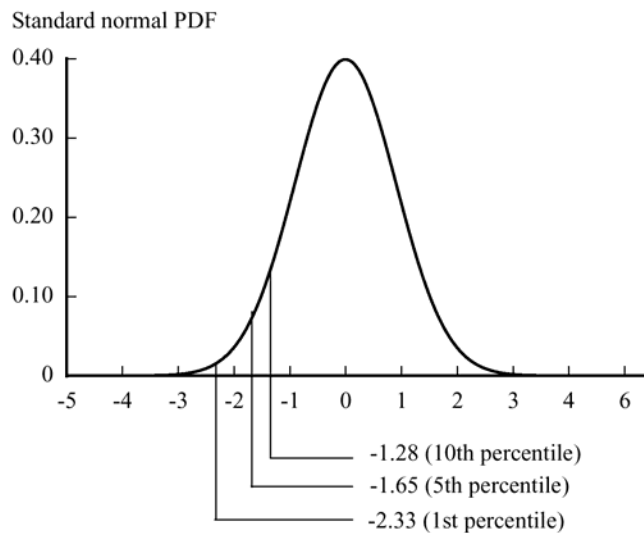


Figure 1: Normal distribution, Source: RiskMetricsTM - Technical Document, p. 69.

1 standard deviation 68.26% probability
2 standard deviations 95.54% probability
3 standard deviations 99.74% probability

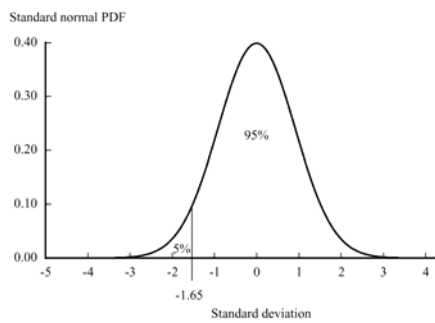


Figure 2: Normal distribution, Source: *CreditMetricsTM-Technical Document, p. 70.*

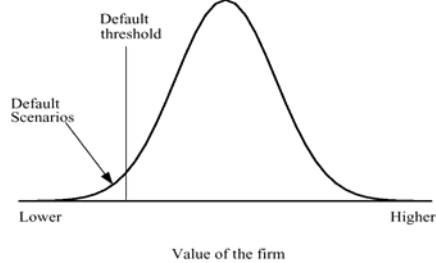


Figure 3: Normal distribution, Source: *CreditMetricsTM-Technical Document, p. 37.*

A Digression to History

In his 1900 dissertation on "The Theory of Speculation," Louis Bachelier searched for "a formula which expresses the likelihood of a market fluctuation. " He ended up with a mathematical formula that describes the Brownian Motion.

In the finance world, Brownian Motion came to be called the random walk, once described as the path a drunk might follow at night in the light of a lamp post.

Using the geometric Brownian motion to describe the random fluctuations in stock prices, Fisher Black, Myron Scholes, and Bob Merton worked out the Black Scholes option pricing formula.

This work was done in the spring of 1970, when both Merton and Scholes were at MIT Sloan!

Model of the Behavior of Stock Prices

Wiener Processes

The change Δz during a small period of time Δt is:

$$\Delta z = \varepsilon \sqrt{\Delta t} \quad (1)$$

where ε is a random drawing from a standardized normal distribution $N(0,1)$. The values of Δz for any two different short intervals of time Δt are independent.

It follows from the first property that Δz itself has a normal distribution with:

$$\text{mean of } \Delta z = 0 \quad (2)$$

$$\text{standard deviation of } \Delta z = \sqrt{\Delta t} \quad (3)$$

$$\text{variance of } \Delta z = \Delta t \quad (4)$$

The second property implies that z follows a Markov process.¹

Consider the increase in the value of z during a relatively long period of time T . This can be denoted by $z(T) - z(0)$. It can be regarded as the sum of the increases in z in N small time intervals of length Δt , where

$$N = \frac{T}{\Delta t} \quad (5)$$

Thus

$$z(T) - z(0) = \sum_{i=1}^N \varepsilon_i \sqrt{\Delta t} \quad (6)$$

where the ε_i ($1, 2, \dots, N$) are random drawing from $N(0,1)$.

¹A Markov process is a particular type of stochastic process where only the present value for a variable is relevant for predicting the future. The past history of the variable and the way that the present has emerged from the past are irrelevant. Stock prices are usually assumed to follow a Markov process. Suppose that the price of IBM stock is \$ 100 now. If the stock price follows a Markov process, our predictions should be unaffected by the price one week ago, one month ago, or one year ago. The only relevant piece of information is that the price is now \$ 100.

Predictions for the future are uncertain and must be expressed in terms of probability distributions. The Markov property implies that the probability distribution of the price at any particular future time is not dependent on the particular path followed by the price in the past.

The Markov property of stock prices is consistent with the *weak form of market efficiency*. This states that the present price of a stock impounds all the information contained in a record of past prices. If the weak form of market efficiency were not true, technical analysts could make above-average returns by interpreting charts of the past history of stock prices. There is very little evidence that they are, in fact, able to do this.

Statistical properties of the stock price history of IBM may be useful in determining the characteristics of the stochastic process followed by the stock price (e.g. its volatility). The point being made here is that the particular path followed by the stock in the past is irrelevant.

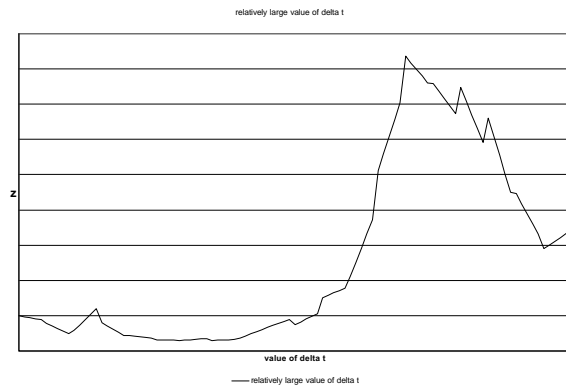


Figure 3: Relatively large value of Δt

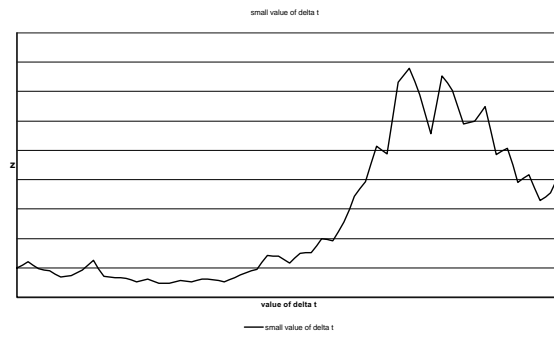


Figure 4: Small value of Δt

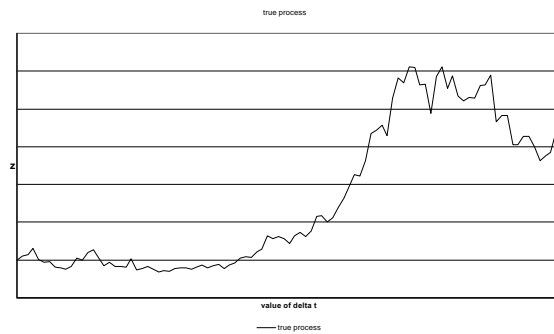


Figure 5: The true process obtained as $\Delta t \rightarrow 0$

From the second property of Wiener processes, the ε_i 's are independent of each other, if follows from equation 6 that $z(T) - z(0)$ is normally distributed with

$$\text{mean of } [z(T) - z(0)] = 0 \quad (7)$$

$$\text{standard deviation of } [z(T) - z(0)] = \sqrt{\Delta t} \quad (8)$$

$$\text{variance of } [z(T) - z(0)] = \Delta t \quad (9)$$

This is consistent with the discussion earlier in this section.

Generalized Wiener Process

The basic Wiener process, dz , that has been developed so far has a drift rate of zero and a variance rate of 1.0. The drift rate of zero means that the expected value of z at any future time is equal to its current value. the variance rate of 1.0 means that the variance of the change in z in a time interval of length T equals T . A generalized Wiener process for a variable x can be defined in terms of dz as follows

$$dx = a dt + b dz \quad (10)$$

where a and b are constants.

To understand equation 10, it is useful to consider the two components on the right-hand side separately. The $a dt$ term implies that x has an expected drift rate of a per unit of time. With the $b dz$ term, the equation is

$$dx = a dt \quad (11)$$

which implies that

$$\frac{dx}{dt} = a \quad (12)$$

or

$$x = x_0 + a t \quad (13)$$

where x_0 is the value of x at time zero. In a period of time of length T , x increases by an amount $a T$. The $b dz$ term on the right-handed side of euqation 10 can be regarded as adding noses or variability to the parth followed by x . The amount of this noise or variability is b times a Wiener process. A Wiener process has a standard deviation of 1.0. It follows that b times a Wiener proces has a standard deviation of b . In a small time interval Δt , the change in value of x , Δx , is from equation 1 and 10, given by

$$\Delta x = a\Delta t + \varepsilon\sqrt{\Delta t} \quad (14)$$

where, as before, ε is a random drawing from a standardized normal distribution. Thus Δx has a normal distribution with

$$\text{mean of } \Delta x = a\Delta t \quad (15)$$

$$\text{standard deviation of } \Delta x = b\sqrt{\Delta t} \quad (16)$$

$$\text{variance of } \Delta x = b^2\Delta t \quad (17)$$

Similar arguments to those given for a Wiener process show that the change in the value of x in any time interval T is normally distributed with

$$\text{mean of } \Delta x = aT \quad (18)$$

$$\text{standard deviation of } \Delta x = b\sqrt{T} \quad (19)$$

$$\text{variance of } \Delta x = b^2T \quad (20)$$

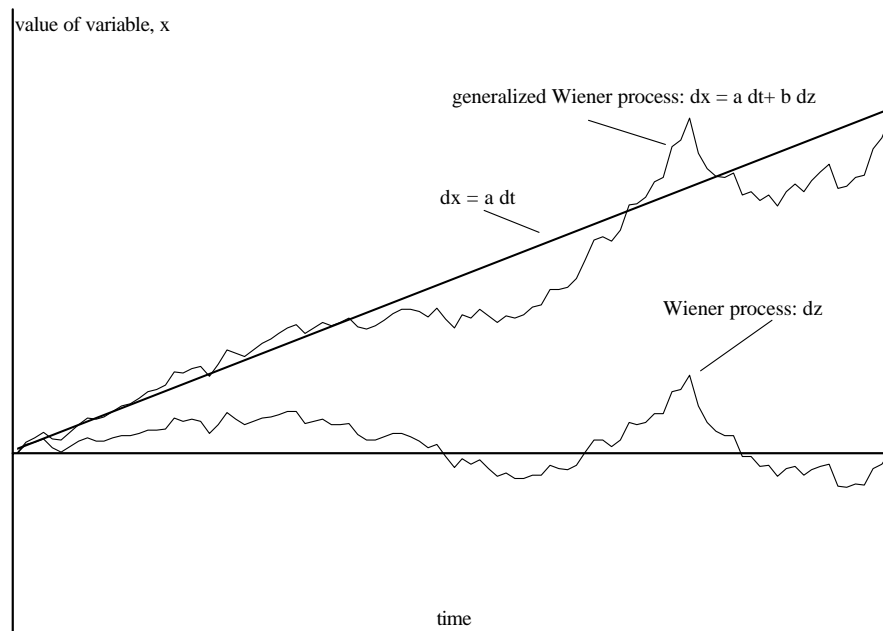


Figure 6: Wiener processes

Thus, the generalized Wiener process given in equation 10 has an expected drift rate (i.e. average drift per unit of time) of a and a variance rate (i.e., variance per unit of time) of b^2 . It is illustrated in Figure (6).

Process for Stock Prices

It is tempting to suggest that a stock price follows a generalized Wiener process; that is, that it has a constant expected drift rate and a constant variance rate. However, this model fails to capture a key aspect of stock prices. This is that the expected percentage return required by investors from a stock is independent of the stock price. If investors require a 14% per annum expected return when the stock price is \$ 10, then *ceteris paribus*, they will also require a 14% per annum expected return when it is \$ 50. Clearly, the constant expected drift-rate assumption is inappropriate and needs to be replaced by the assumption that the expected return (that is, expected drift divided by the stock price) is constant. If S is the stock price at time t , the expected drift rate in S should be assumed to be μS for some constant parameter, μ . This means that in a short interval of time, Δt , the expected increase in S is $\mu S \Delta t$. The parameter, μ , is the expected rate of return on the stock, expressed in decimal form.

If the volatility of the stock price is always zero, this model implies that

$$\Delta S = \mu S \Delta t \tag{21}$$

in the limit as $\Delta t \rightarrow 0$

$$dS = \mu S dt \tag{22}$$

or

$$\frac{dS}{S} = \mu dt \tag{23}$$

so that

$$S_T = S_0 e^{\mu T} \tag{24}$$

where S_0 and S_T are the stock price at time zero and time T . Equation 24 shows that when the variance rate is zero, the stock price grows at a continuously compounded rate of μ per unit of time.

In practice, of course, a stock price does exhibit volatility. A reasonable assumption is that the variability of the percentage return in a short period of time, Δt , is the same regardless of the stock price. In other words, an investor is just as uncertain of the percentage return when the stock price is \$ 50 as when it is \$ 10. This suggests that the standard deviation of the change in a short period of time Δt should be proportional to the stock price and leads to the model

$$dS = \mu S dt + \sigma S dz \tag{25}$$

or

$$dS/S = \mu dt + \sigma dz. \tag{26}$$

Equation 26 is the most widely used model of stock price behavior. The variable σ is the volatility of the stock price. the variable μ is its expected rate of return.

The **discrete-time version** of the model is

$$\frac{\Delta S}{S} = \mu\Delta t + \sigma\varepsilon\sqrt{\Delta t} \quad (27)$$

$$\Delta S = \mu S\Delta t + \sigma S\varepsilon\sqrt{\Delta t} \quad (28)$$

The variable ΔS is the change on the stock price, S , in a small interval of time, Δt , and ε is a random drawing from a standardized normal distribution (i.e.g, normal distribution with a mean of zero and standard deviation of 1.0). The parameter, μ , is the expected rate of return per unit of time from the stock and the parameter, σ , is the volatility of the stock price. Both of these parameters are assumed constant.

The left-hand side of equation 27 is the return provided by the stock in a short period of time, Δt . The term $\mu\Delta t$ is the expected value of this return, and the term $\sigma\varepsilon\sqrt{\Delta t}$ is the stochastic component of the return. The variance of the stochastic component (and therefore, of the whole return) is $\sigma^2\Delta t$. This is consistent with the definition of the volatility, σ . That is, σ is such that $\sigma\sqrt{\Delta t}$ is the standard deviation of the return in a short time period, Δt .

Equation 27 shows that $\Delta S/S$ is normally distributed with mean $\mu\Delta t$ and standard deviation $\sigma\sqrt{\Delta t}$. In other words,

$$\frac{\Delta S}{S} \approx N\left(\mu\Delta t, \sigma\sqrt{\Delta t}\right) \quad (29)$$

Why Normals?

Model the random fluctuation of stock prices using geometric Brownian motion. Implication for stock returns: normal distribution (for continuously compounded returns).

The annualized stock return is normal with mean (μ) and standard deviation (σ) . For the S&P 500 index returns, μ is roughly 12%, and σ is roughly 15%. The latter is also called volatility.

Fix a time horizon, say Δt . The stock return over Δ is normally distributed with mean $\mu\Delta t$ and standard deviation $\sigma\sqrt{\Delta t}$. What is the distribution of daily returns?

Events that are not Normal

A negative surprise: on October 19, 1987, the S&P 500 index dropped more than 23% on one day.

A positive surprise: on January 3, 2001, the Nasdaq composite index gained more than 14% on one day.

Suppose we use normal distribution to characterize daily stock returns. What are the probabilities of such surprises?

A Positive Surprise

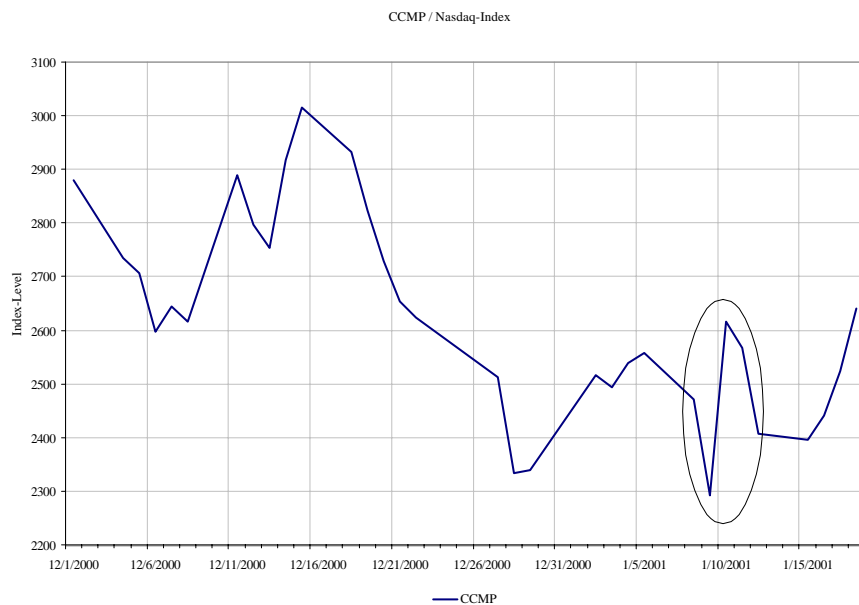


Figure 4: Nasdaq Dec 1 to Jan 18 2001, jumps on January 2nd and 3rd

The Probability of a Crash

Let r denote the daily return, which is:

- normally distributed with
- mean $0.12/252 = 0.00048$,
- standard deviation $0.15 \cdot \sqrt{252} = 0.0094$

What is the probability of an '87 crash? $\text{Prob}(r < 0.23) = ?$

- First, convert r to a standard normal

$$X = \frac{r - 0.00048}{0.0094} \quad (30)$$

- Second, convert the critical value for r to that for X :

$$\frac{-0.23 - 0.00048}{0.0094} \approx -23 \quad (31)$$

- Finally, knowing that X is a standard normal,

$$\text{Prob}(X < -23) = N(-23) = 10^{-127!} \quad (32)$$

What Normal Fails to Capture...

There are large movements (both up and down) in stock prices that cannot be captured at all by the normal distribution.

In mathematical terms, the tail distribution of a normal random variable is too thin. Historical stock returns exhibit fat tails.

If we make financial decisions based on normal distribution, we will miss out on the large movements. The consequences are catastrophic!

This is especially important for leveraged investments over a short time horizon.

Tail fatness is also an important issue in risk management.

Data Analysis

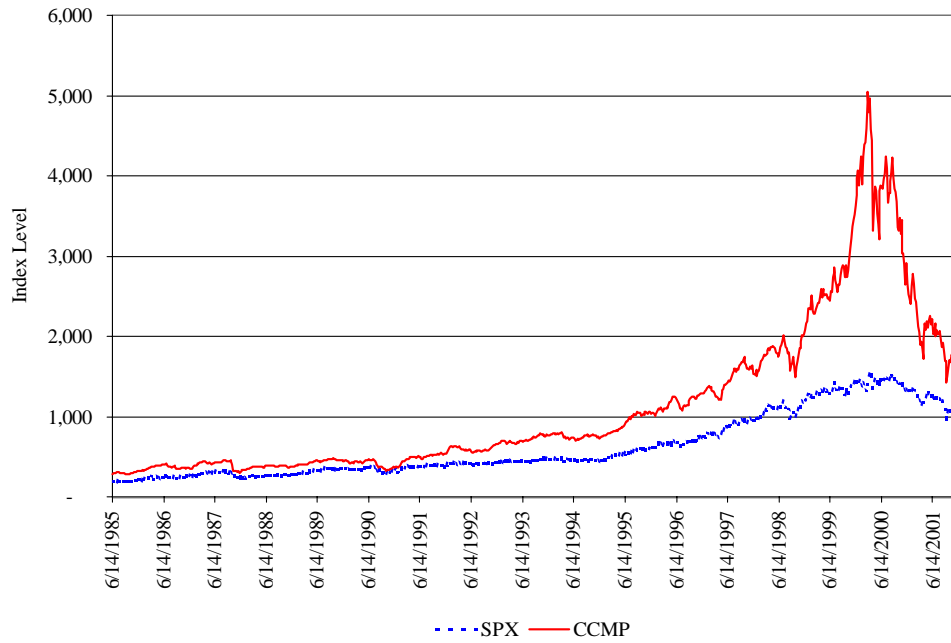


Figure 7: SP 500 and Nasdaq Index, index points, source: Bloomberg Professional.

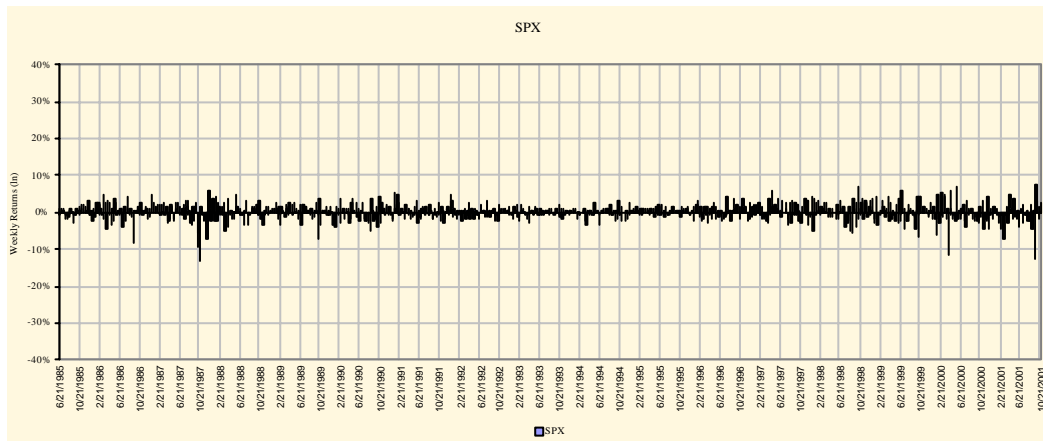
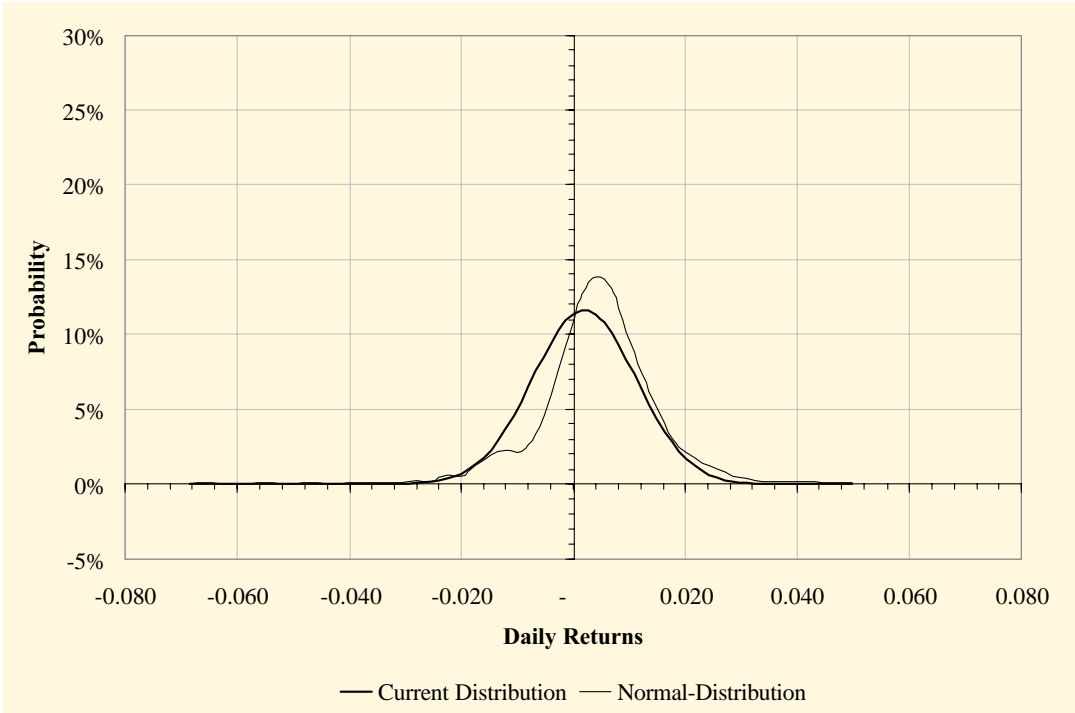


Figure 8: SP 500 daily returns, source: Bloomberg Professional



Preliminaries for Data Analysis

When given the "raw data," first look for trends. If there are any, the first step is always to **de-trend** the data.

Why?

The i.i.d. assumption for r_1, r_2, \dots, r_N : returns are independent and identically distributed.

The longer we observe, the more we know about the probability distribution . . . but do not forget **structural changes!**

Empirical Distribution

1. Sort the outcomes r_1, r_2, \dots, r_N
 2. Denote the minimum by \underline{x} and the maximum by \bar{x} . Divide $[\underline{x}, \bar{x}]$ evenly into K bins:
bin 1: $[\underline{x}, \underline{x} + \Delta x]$
bin 2: $[\underline{x} + \Delta x, \underline{x} + 2\Delta]$
...
bin 3: $[\bar{x} - \Delta x, \bar{x}]$
- where $\Delta x = (\bar{x} - \underline{x}) / K$
3. Fixing a number K , count the number N_k of the r_i 's that fall to the k -th bin.
 4. Repeat this for $k = 1, 2, \dots, K$, we get a list of bins and their respective "likelihood" N_k .
 5. Finally, we need to renormalize the likelihood so that the probability of r_i falling into $[\underline{x}, \bar{x}]$ is one.

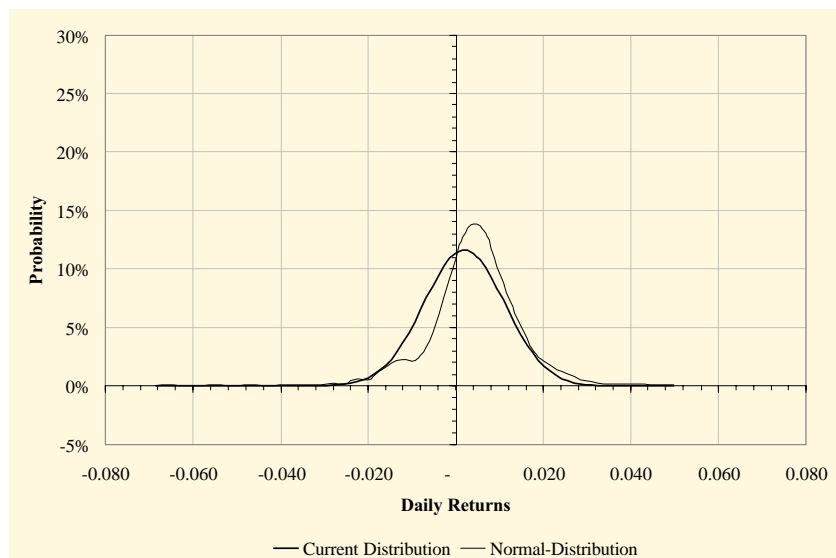


Figure 10: Daily return distribution of S&P 500 index.

source: Bloomberg Professional

Fat Tails

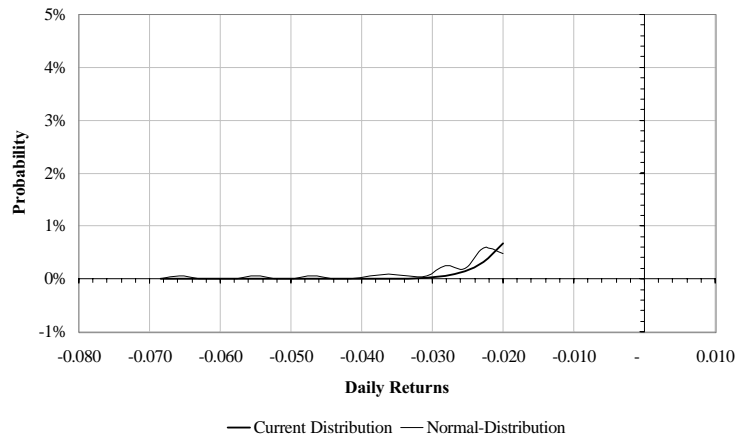


Figure 11: Left fat tail of daily Return distribution of S&P 500 index, 5% on left hand side.

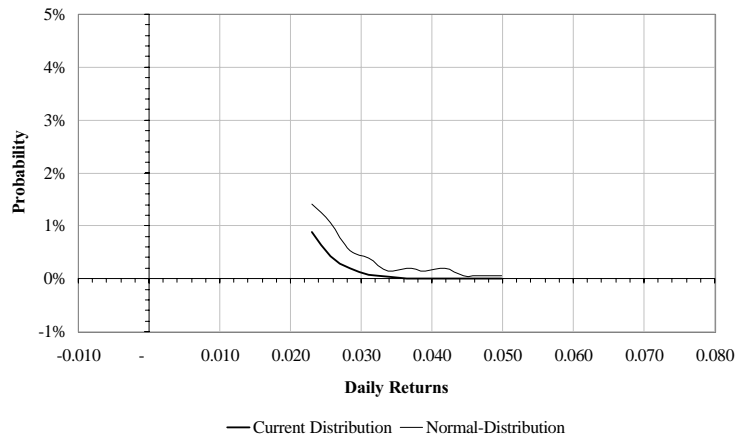


Figure 12: Right fat tail of daily Return distribution of S&P 500 index, 5% on right hand side

Sample Statistics

mean

$$\mu = \frac{1}{N} \sum_{i=1}^N r_i \quad (33)$$

variance:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (r_i - \mu)^2 \quad (34)$$

skewness (lack of symmetry):

$$skew = \frac{\frac{1}{N} \sum_{i=1}^N (r_i - \mu)^3}{\sqrt{\sigma^3}} \quad (35)$$

kurtosis (peakedness):

$$kurt = \frac{\frac{1}{N} \sum_{i=1}^N (r_i - \mu)^4}{\sigma^2} \quad (36)$$

Standard Errors

Take the sample mean μ as an example:

$$\text{mean: } \mu = \frac{1}{N} \sum_{i=1}^N r_i$$

We know that r_i 's are random draws from a stationary (and ergodic²) distribution. In fact, to simplify our analysis, we assumed that they are i.i.d.

This implies that the sample mean μ is itself a random variable.

- What is its mean?
- What is its standard deviation?

Standard errors: measure the precision of the estimators.

²ergodic: Of, pertaining to, or possessing the property that in the limit all points in a space are covered with equal frequency, or that each sufficiently large selection of points is equally representative of the whole. [Oxford English Dictionary]

Conditional Version

So far, we assumed that the distribution of stock returns stays the same throughout history. We calculate the sample statistics as if history repeats itself with the same probability law.

But we know that this cannot be true. How do we use the data when we have reasons to believe that the probability law changes over time?

For example, suppose we believe that the daily returns within each month t are normally distributed with mean μ_t and standard deviation σ_t , which vary over time t .

The simplest way to take the conditional information into account is to calculate the sample mean and standard deviation month by month.

Time Series Patterns

While the unconditional version of sample statistics gives us a static picture of daily stock returns, the conditional version provides more dynamic information. For example,

- The conditional expected returns are time varying, but there is not much persistence.
- The conditional volatilities are also time varying. Moreover, they seem to be highly persistent.
- There is a negative relationship between returns and volatility: when the market goes down, the volatility goes up.

We will revisit these issues in more detail in Class 9.

Summary

We use probability distributions to characterize and evaluate random events.

There is a long tradition of using normal distributions to characterize the fluctuations in stock prices.

The normal distributions, however, are not adequate to capture large surprises.

The empirical distribution and sample moments are useful statistical tools to extract information from the data. Stationarity is one important assumption.

The precision of the sample moments can be measured by their standard errors.

The sample statistics can be used in both unconditional and conditional versions. The conditional version provides more dynamic information about the data.

Focus:

BKM Chapters 3 & 5 (Chapter 3: This is all general stuff, you need to know the basic things like IPO, private placements, secondary market etc.);

- p. 137 (probability distribution, standard deviation);
- p. 141 (figure 5.4);
- p.149 bottom and 150 (continuous compounding);

Reader: Fama (1995).

Type of potential questions: chapter 3 concept check question 2-3, p. 98 ff. question 2, 5, 11, 17, 22 chapter 5, p. 146 ff. question 10, 12, 15

Questions for Next Class

Please read:

- BKM Chapters 6 and 7,
- Elton and Gruber (2000), and
- Kritzman (1992)

Think about the following questions:

- Two important components in making an investment decision: the investment opportunity and the investor.
- Admittedly, one investor may differ from another. And the investment opportunity does not stay constant across space or time.
- If you were asked to build a model of investments for a generic investor on a generic market environment, what are the basic features you would include in your model?