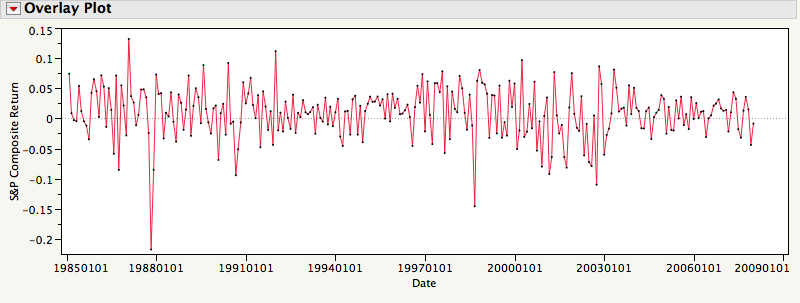
**Module 2: Probability Models**

# Application: Pricing a Digital Option

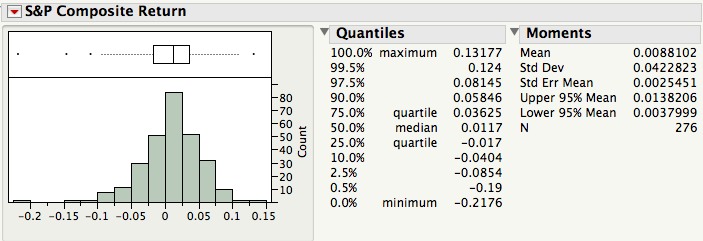
Many investors place some of their assets in stocks, such as purchasing an index fund.

The history of the stock market reveals that investors face risks. Although the market has headed up for most of its history, some months have shown very large declines. The data file *total\_market.JMP* contains monthly returns on the composite S&P index.

This plot shows these returns from 1985 through 2007. Do you recognize any dates?



We can summarize these data with a histogram that shows the distribution of the monthly returns over these years.



The average monthly return is 0.88% with standard deviation 4.23%. What do these suggest?

Investors with large positions in the stock market might want to avoid some of these risks. Rather than selling their stocks and paying all of those fees, investors can purchase a type of insurance in the form of an *option*.

We’ll consider a very special type of option known as a *digital option*. This can be used to construct more complicated derivatives; an asset, such as an option that *derives* its value from another asset, is known as a derivative.

The question of interest:

To insure yourself against a large drop in the Standard & Poors (S&P) index, how much should you pay for a digital option that pays $1 if the S&P index drops by more than 10% next month?

The price for one option is not so interesting. But consider an investor who wants to buy 1,000,000 of these options.

What’s a good price for the option from the point of view of the investor? From the point of view of the bank that offers the option?

The complicating factor:

We need to anticipate what might happen in a *future month*. All we have to go on is what has happened in the past. An abstraction that we’ll use to think carefully about such problems is known as a random variable.

# Random Variables

A *random variable* (rv) (e.g., X, Y, Z, etc.) is used to represent an uncertain quantity. (You can think of it as the uncertain numerical outcome of an experiment or process that has yet to occur).

A random variable associates a number between 0 and 1 with each possible outcome of an experiment or random process. This number is called the *probability* of the outcome.

Useful Notation: If X is a random variable, then

P(X = x) is the probability that X will take on the value x.

P(x1 ≤ X ≤ x2) is the probability that X will take on a value in the interval [x1, x2].

The Long Run Manifestation of Probability: intuitively, P(X = x) is the proportion of times

X = x over an “infinitely” long sequence of repetitions of the experiment.[[1]](#footnote-1)

In this sense, a graph of P(X = x) is the histogram of outcomes in this long sequence.

# Example: Toss a Fair Coin

Random variable: X Equally likely possible outcomes: 1 if heads, 0 if tails

Probabilities: P(X = 1) = ½, P(X = 0) = ½

# Example: Play Roulette

38 slots labeled 00, 0, 1, 2, …, 36 Equally likely possible outcomes

1-36 colored Red (if even) and Black (if odd)

For $1, place a bet on Red

Pays you $2 if randomly selected slot is Red

Random variable: R = Net winnings Possible outcomes: -1 and 1

Probabilities: P(R = -1) = 20/38, P(R = 1) = 18/38

# Example: Play the Lottery

For $1, buy a lottery ticket numbered from 000 to 999.

Pays you $500 if ticket matches randomly selected number.

Not equally likely!

Random variable: D = Net winnings Possible outcomes: -1 and 499

Probabilities: P(D = -1) = 999/1000 P(D = 499) = 1/1000

# Example: Pick a Chip

Randomly draw a chip from a bowl containing 10,000 chips with these values:

5,000 $1 chips

3,000 $5 chips

1,000 $10 chips

1,000 $20 chips

Random variable: X = value of the drawn chip Possible Outcomes: 1, 5, 10, and 20

Probabilities: P(X = 1) = 5/10, P(X = 5) = 3/10, P(X = 10) = 1/10, and P(X = 20) = 1/10

For convenience, we often write p(x) = P(X = x), so that

p(1) = 5/10 p(5) = 3/10 p(10) = 1/10 p(20) = 1/10

The function p(x) is called the *probability distribution* of X.

Note that:

p(x) = 0 for x ≠ 1,5,10,20

p(1) + p(5) + p(10) + p(20) = 1

The probability of any *event* concerning X can be computed from p(x). For example:

P(X ≤ 5) = 5/10 + 3/10 = 8/10

P(5 ≤ X ≤ 10) = 3/10 + 1/10 = 4/10

P(X is an even number) = 1/10 + 1/10 = 2/10

Let’s draw a histogram of the population of chips in the bowl:

How does this shape compare to the shape of the probability distribution?

This bowl of chips can be thought of as a population from which the chip was drawn.

Useful interpretation: the outcome of random variable with probability distribution p(x) *can always be interpreted as* a random selection from a p(x) shaped population distribution.

Let’s now compute the mean of the chip histogram above.

# The Expected Value of a Random Variable

Suppose the N possible outcomes of a rv X are:

x1, x2,…, xN.

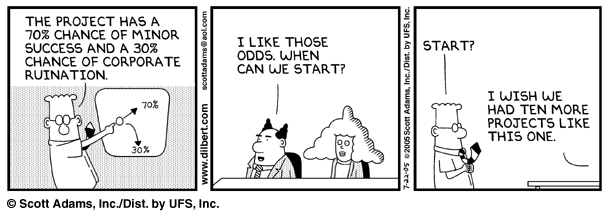
The *expected value* of X, denoted by E[X], is defined as a weighted sum of the possible outcomes of X[[2]](#footnote-2):

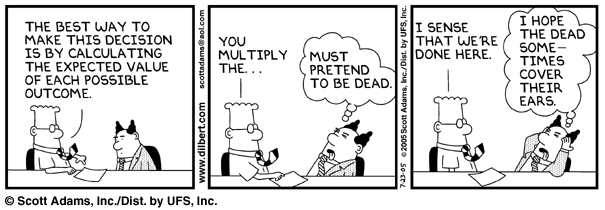
E[X] = x1 p(x1) + x2 p(x2) + … + xN p(xN) = 

Sometimes the expected value of X is called the *mean* of X and is denoted by the Greek letter *μ*. This is the same as the average value in the population from which X is drawn. For example, for the chips we find the average value is:



This is the population version of the average of a collection of numbers.



[[3]](#footnote-3)

# Examples

Example: R = Winnings in Roulette

E[R] = -1 p(-1) + 1 p(1) = -20/38 + 18/38 = -1/19

Example: D = Winnings in Daily Numbers Lottery

E[D] = -1 p(-1) + 499 p(499) = -999/1000 + 499/1000 = -0.5

Example: X = Value of the Drawn Chip

μ = 1 p(1) + 5 p(5) + 10 p(10) + 20 p(20) = 5/10 + 15/10 + 10/10 + 20/10 = 5

# Interpretations of the Expected Value or Mean

Probability-weighted average of possible outcomes

Center of the probability distribution of X

Mean of the population from which the rv is drawn

Long-run average over “infinitely” many repetitions of X

“Fair value” of a gamble

# Jargon

μ is sometimes called the *population mean*

to distinguish it from , the *sample mean.*

# Variance and Standard Deviation of a Random Variable

As with our descriptions of a histogram, it’s not enough to only know the mean of a random variable. We also need a summary of its variation.

Suppose the possible outcomes of a rv X are

x1, x2,…,xN.

Then the *variance* of X, denoted Var[X] or σ2, is

Var[X] = σ2  = (x1 − μ)2 p(x1) + … + (xN − μ)2 p(xN) = 

An equivalent definition expressed using an expected value is

Var[X] = σ2 = E[(X− μ)2 ]

The *standard deviation* of X, denoted SD[X] or σ, is the square root of the variance.

# Examples

Example: R = Net Winnings in Roulette

Var[R] = (-1 − (-1/19))2 p(-1) + (1 − (-1/19))2 p(1) = 0.9972

SD[R] = 0.9986

Example: D = Net Winnings in Daily Numbers Lottery

Var[D] = (-1 − (-0.5))2 p(-1) + (499 − (-0.5))2 p(499) = 249.75

SD[D] = 15.80

Example: X = Value of the Drawn Chip

σ2  = (1 − 5)2 p(1) + (5 − 5)2 p(5) + (10 − 5)2 p(10) + (20 − 5)2 p(20) = 33

σ = 5.744

# Interpretations of the Variance

Probability-weighted average squared deviation

Dispersion of the probability distribution of X

Variance of the population from which the r.v. is drawn

Long run average squared deviation over “infinitely” many repetitions of X

“Risk” of a gamble (more coming in Module 4)

# Jargon

σ2 and σ are sometimes called the *population variance* and the *population standard deviation*

to distinguish them from s2 and s, the *sample variance* and the *sample standard deviation*.

# Continuous Random Variables

As opposed to the random variables we have considered, some random variables represent draws from a population of values so dense that it is easier to model the shape of the population by a smooth continuous function.

# Example: A Normal Random Variable

X = value drawn randomly from a normal population with mean μ and standard dev σ

Often abbreviated as X ~ N(μ, σ2)

Population shape: 



Useful Interpretation: For X, a continuous random variable, P(x1 ≤ X ≤ x2) is given by the area under the continuous curve between x1 and x2.

Some familiar probabilities:

P(μ − σ ≤ X ≤ μ + σ) ≈ 68%

P(μ − 2σ ≤ X ≤ μ + 2σ) ≈ 95%

P(μ − σ ≤ X ≤ μ + 2σ) ≈ 81.5%

P(−∞ < X < ∞) = 1

More generally, normal probabilities for any given values of μ and σ can be computed using the Normal Distribution probability formula in JMP.

For example, when X ~ N(0.5, 1.82),

P(X ≤ 0.8) = 0.566 and P(0.8 ≤ X ≤ 1.4) = 0.125.[[4]](#footnote-4)

# Application: Pricing a Digital Option

Let X = the return on the S&P index next month

Suppose you had strong enough evidence to *assume* that X was a normal random variable with mean .0088 and standard deviation .0423.

These values for the mean and standard deviation match the month-to-month performance of the stock market from January 1985 to December 2007.

Let’s return to our question on pg 2-3: how much should you pay for a digital option that pays $1 if the S&P drops by more than 10% next month?

Random variables provide a precise way to find the value of the option. We can express the value of such an option by defining a new, discrete random variable Y whose values are determined by X (*i.e*., a derivative since its value is derived from that of another asset), with

Y = $1 if X < –.10

= $0 otherwise

A fair price for the option would then be E[Y] = P[Y = 1] = P[X < –.10] = .0051 [[5]](#footnote-5)

# The Normal Distribution as a Model

The value of E[Y] = P[X < –.10] = .0051 depends strongly on the assumption that X would be a draw from a particular normal distribution.

An outline of the steps used to make such an assumption:

1. Verify that the behavior of the time series of past monthly S&P returns is consistent with assuming a sequence of independent draws from the same population.
2. Verify that the histogram and normal quantile plot of the past monthly returns are consistent with assuming a normal population.
3. Estimate μ and σ2 of the population by the values of  and s2 for the past monthly returns.

An alternative “common sense” procedure might be to estimate P[X < –.10] directly from the data: use the proportion of prior months that the S&P fell below –.10. What is the drawback of such an approach?

Compared to this simplistic procedure, the normal assumption allows us to use the data to extrapolate more efficiently.[[6]](#footnote-6)

# The (Student) t Distributions

The t distributions are a family of continuous probability distributions that resemble normal distributions.

They are symmetric and bell-shaped, but have heavier or fatter tails.

Because of their heavier tails, extreme observations are more likely with draws from a t distribution.



Each t distribution is identified by 3 parameters: μ, σ2, and a shape parameter called the degrees of freedom (df)**.** df is always positive. As df gets larger, *t* distributions more closely approximate normal distributions.[[7]](#footnote-7) *t* distribution probabilities can be obtained by the probability function “t distribution” on the JMP calculator.[[8]](#footnote-8)

If we treat the return on the S&P index next month X as having a t-distribution that is matched to the market during the 1985-2007 period, we obtain:

P(X < –.10)  .03

Why does this probability differ so muchfrom the previous normal probability calculation? (It’s about 6 times larger.)

Quantile plots can also be used as a diagnostic check for t distribution assumptions.[[9]](#footnote-9)

Key point:

In this example, we have two different models for the data. Both describe the data reasonably well, so that the data are not clear about which model is better.

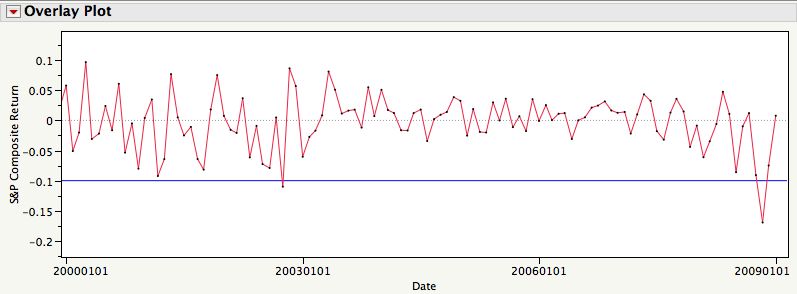
The model matters: different models produce different prices for the options.

In cases like this, we need to be sensitive to the consequences of our assumptions. We may never know which model is “right,” but we should be aware of the alternatives.

**What Happened in 2008**

We have hidden these rows, but the data file includes 2008.

The S&P Index fell by more than 10% during one month of 2008, and came very close to this threshold in another!



# Take-Away Review

Random variables provide a useful notation for describing problems that have uncertain outcomes.

The outcomes of a random variable can be associated with probabilities.

The expected value is the average of the outcomes, weighted by probabilities.

The variance is the probability weighted average squared deviation.

If we match a random variable to features of a sample, we can use that random variable to model the underlying process.

But, the results we obtain are sensitive to the model that we use.

Check the assumptions of your model.

# Next Module

Random variables and returns on investments.

1. See section 7.1 in SF, page 151, for more information on the relationship between data and probabilities. Random variables are introduced in Chapter 9. [↑](#footnote-ref-1)
2. SF, page 200. [↑](#footnote-ref-2)
3. Alas, Dilbert is not a statistician. You want the expected value of a project, not each outcome! [↑](#footnote-ref-3)
4. In JMP, the probability P(X < a) when X ~ N(μ, σ2) is obtained with the JMP formula Normal Distribution(a, μ, σ) which can be obtained as one of the Probability functions in the formula window. Note that the entries for μ and σ become available by inserting commas after the first entry. The calculation here is illustrated in the file *Norm Prob.JMP*. [↑](#footnote-ref-4)
5. The calculation of this probability is illustrated in the file *Option Prob.JMP*. In Excel, use the function NORMDIST. [↑](#footnote-ref-5)
6. Be careful, however! Other models may give rather different prices. Although normal distributions approximate many distributions that naturally occur, important alternatives may also be considered. Some are hard to tell from the normal. [↑](#footnote-ref-6)
7. Figures 15.3 and 15.4 in SF, page 358, compare the shapes of the normal and t distributions. [↑](#footnote-ref-7)
8. You can use tables like the one in SF (inside back cover), but it is much easier and more accurate to let software do these calculations. [↑](#footnote-ref-8)
9. JMP does not implement these alternative quantile plots. The plots would look like normal quantile plots, but the reference model would a t-distribution rather than a normal model. [↑](#footnote-ref-9)