Stat 107: Introduction to Business and Financial Statistics

Class 4: Statistics Review, Part I

FYI-Introduction to Investments

https://www.ocf.berkeley.edu/~jml/decal/download.html

Yuck

Hyperbole Alert On The Street - Federal Reserve One Of The Best Central Banks ...

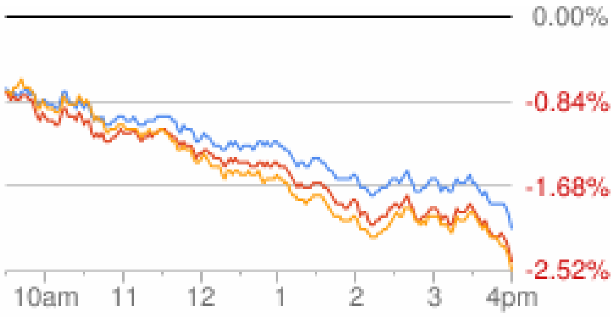
Forbes - 3 hours ago

Yes, OK, we know that certain parts of the financial press like to report in a rather less restrained manner than some others of us do but it is possible to go a little too far with this.

Worst Central Bank? That's our Fed   TheStreet.com

Futures File: Oil swings on supply shocks   Casper Star-Tribune Online

USA TODAY   Seeking Alpha



Dow Jones	18,085.45	-394.46	(-2.13%)
S&P 500	2,127.81	-53.49	(-2.45%)
Nasdaq	5,125.91	-133.57	(-2.54%)

3,511 of 4,929 stocks are above their 200 DMA (71.23%)

Currently 71.23% of stocks tracked by InvestProfits are trading above their 200 Day Moving Average (DMA). That means 3,511 of the 4,929 stocks are trading above their 200 day moving average. Above 80 percent is bearish; 20 percent and below is bullish. To filter out the volatility of low priced stocks, those selling below a price of $5 per share were omitted.

Theory Week

- This section of the course reviews the probability concepts that are necessary for the modeling and statistical analysis of financial data.
- This material is typically covered in an introductory probability and statistics course.
- In the course of the review, some examples related to finance will be presented and some important risk concepts, such as value-at-risk, will be introduced.
- The examples will also show how to use R for probability calculations and working with probability distributions.

Theory Week

- Monday
  - Random Variables
  - E(X), Var(X)
  - Cov(X,Y)
  - E(X+Y) Var(X+Y)
  - Binomial, Normal
- Wednesday
  - CLT
  - Confidence Intervals
  - Hypothesis Testing

Univariate Random Variables

- A random variable (rv) *X* is a variable that can take on different values with different probabilities.
- Random variables can be discrete or continuous.

Examples

- X=price of $MSFT next month
- X=return on a one month investment
- X=1 if stock price goes up; X=0 if stock price goes down.

Discrete Random Variables

- We assign probabilities to all the values that the random variable can take on. There are certain requirements about the values we assign.
- If a random variable X can take values x_i, then the following must be true:

$$(1)\ 0 \le P_X(x_i) \le 1$$

$$(2)\ \sum_{all\ x_i} P_X(x_i) = 1$$

$P_X(x)$ is sometimes called the probability distribution function

Example

- Probability distribution for yearly return of $MSFT

Return	Prob
-0.3	0.05
0	0.2
0.1	0.5
0.2	0.2
0.5	0.05

Expected Value

- For a random variable, the analogy to the sample mean is called the expectation or expected value. The letter E usually denotes an expected value, and this symbol is usually followed by brackets enclosing the random variable of interest.
- Definition: Given a random variable X with values x_i, the expected value of X is

$$\mu_X = E(X) = \sum_{all\ x_i} x_i P(X = x_i)$$

$$= \sum_{all\ x_i} x_i P(x_i)$$

Example

- For the MSTF example

Return	Prob
-0.3	0.05
0	0.2
0.1	0.5
0.2	0.2
0.5	0.05

$$\begin{aligned} E[X] &= (-0.3)\cdot(0.05) + (0.0)\cdot(0.20) + (0.1)\cdot(0.5) \\ &\quad + (0.2)\cdot(0.2) + (0.5)\cdot(0.05) \\ &= 0.10. \end{aligned}$$

Some Expectation Rules

- $E(cX) = cE(X)$ where c is any constant
- $E(X+c) = E(X)+c$
- $E(X+Y) = E(X)+E(Y)$ for any two random variables
- Note-in general combining random variables can lead to some issues, but expectations always behave nicely.

Dropping Coins

- Suppose a game involves dropping 3 coins on the table—a nickel, a dime, and a quarter. Each coin that lands "heads up" you are allowed to keep, so that the possible reward R ranges from 0 to 40c.
- Find the mean of R.
- This can be done by brute force by first finding the distribution of R, then the mean.
- But using the idea of sums of random variables, there is an easier approach.

Breaking down the problem

- Define the following random variables
  - X1 = the nickel's contribution to the reward
  - X2 = the dime's contribution to the reward
  - X3 = the quarter's contribution to the reward

- Then R = X1 + X2 + X3
- E(X1) = .5(.05)+.5(0) [why??]
- E(X2) = .5(.1)+.5(0)
- E(X3) = .5(.25)+.5(0)
- E(R) = E(X1)+E(X2)+E(X3) = 0.20

A betting example

Find the expected value of each of the following bets:

a) you get \$5 with probability 1.0.
b) you get \$10 with probability 0.5, or \$0 with probability 0.5.
c) you get \$5 with probability 0.5, \$10 with probability 0.25 and \$0 with probability 0.25.
d) you get \$5 with probability 0.5, \$105 with probability 0.25 or lose \$95 with probability 0.25.

Which bet is the best bet ?

Variance of a Random Variable

- The mean of a random variable is a very useful and informative quantity, but we are often interested in other measures of a distribution.
- The variance and standard deviation are measures of the dispersion of a random variable around its mean.

Technically, the variance of a random variable is the *expected value of the squared deviation of a random variable from its mean*. **Phew**.

The general mathematical formula is

$$\sigma_X^2 = Var(X) = E[(X - \mu_X)^2]$$

For discrete random variables, this simplifies to

$$\sigma_X^2 = Var(X) = E[(X - \mu_X)^2] = \sum_{all\ x_i} (x_i - \mu)^2 P(X = x_i)$$

The standard deviation is $\sigma_X = \sqrt{\sigma_X^2}$ of ~~course~~

weighted average of deviations from the mean

Alternative Variance Formula

- It can be shown that another formula for variance is

$$\sigma_X^2 = Var(X) = E[X^2] - (\mu_X)^2$$

- This formula is sometimes easier to compute than the previous variance formula.

MSFT Example

- Calculating the variance requires a bit of work-its not difficult but be careful with the math:

Return	Prob
-0.3	0.05
0	0.2
0.1	0.5
0.2	0.2
0.5	0.05

$$\begin{aligned}\mathrm{Var}(X) &= (-0.3-0.1)^2\cdot(0.05)+(0.0-0.1)^2\cdot(0.20)\\ &\quad+(0.1-0.1)^2\cdot(0.5)+(0.2-0.1)^2\cdot(0.2)\\ &\quad+(0.5-0.1)^2\cdot(0.05)\\ &= 0.020\end{aligned}$$

$$\mathrm{SD}(X)=\sigma_X=\sqrt{0.020}=0.141.$$

Basic Rules for Variance

- $Var(X+c) = Var(X)$ for any constant c
- $Var(cX)=c^2Var(X)$
- In general, $Var(X+Y) \neq Var(X)+Var(Y)$ [more on this later]

Interpreting the Variance

- The standard deviation of a random variable is important.
- W e already know that the random variable X can assume different values $x_1, x_2, \ldots, x_n$.
- The expected value of X gives the central location, but it is quite possible that the actual outcome will differ from E(X).
- However, Chebyshev's rule also holds for random variables, so we know that at least 75% of the time, the random variable outcome will be within 2 standard deviations of the mean value.
- More info on interpreting in a few classes from now.

Chebyshev's rule for random vars.

	Chebyshev's Rule	Empirical Rule
	Applies to any probability Distribution	Applies to probability Distributions that are mound Shaped and symmetric
$P(\mu-\sigma<x<\mu+\sigma)$	≥ 0	$\approx 68\%$
$P(\mu-2\sigma<x<\mu+2\sigma)$	$\geq 75\%$	$\approx 95\%$
$P(\mu-3\sigma<x<\mu+3\sigma)$	$\geq 89\%$	$\approx 100\%$

Sharpe Measure

- The Sharpe measure:

$$\text{Sharpe measure} = \frac{\mu}{\sigma}$$

$$\begin{aligned} \text{where } \mu &= \text{ average return} \\ \sigma &= \text{ standard deviation of returns} \end{aligned}$$

- The ***Sharpe measure*** evaluates return relative to total risk
- Higher is always better.

People Screen using the Sharpe Ratio

ft.com > markets > marketsdata >

Fund & ETF Screener

Home | UK | World | Companies | Markets | Global Economy
fastFT | Alphaville | FTfm | Markets Data | Trading Room | Equities | Curr

TOP RISK ADJUSTED LARGE CAP VALUE FUNDS

This screen provides a list of US Large Cap Value Funds that have a Sharpe Ratio of greater than 1.

View 1,232 matches | View criteria

SHARE CLASS NAME	INVESTMENT FOCUS	SHARPE RATIO
Advance Capital I Core Equity Fund Institutional Share	Large Value	1.66%
Advance Capital I Core Equity Fund Retail Share	Large Value	1.64%
Al Frank Fund Class Advisor	Large Value	1.65%
Al Frank Dividend Value Fund Class Advisor	Large Value	1.61%

A Sharpe Ratio Problem

- An old unsolved problem with the Sharpe ratio has been that when its numerator is negative, it gives incorrect ranking of portfolios!

A Sharpe Ratio Problem

- For example, consider two portfolios with a common mean return of -1, and distinct standard deviations of 1 and 10.
- Then the Sharpe ratios are -1 and -0.1, implying that portfolio 2 is more desirable.

A Sharpe Ratio Problem

- Taking absolute values doesn’t help.
- Consider two portfolios:
  - Portfolio A: mean = -1, std dev = 1
  - Portfolio B: mean = -3, std dev = 1
- Then absolute values would imply the second portfolio is better though its not.

Rules for Expectation and Variance

Let X be a random variable with mean μ and variance σ^2. Let a and b be any constant fixed numbers. Define the random variable W=a+bX.

Then

$$\mu_W = E(W) = E(a + bX) = a + bE(X) = a + b\mu_X$$

$$Var(W) = Var(a + bX) = b^2\sigma_X^2$$

The Most Common Linear Transformation

- The Z-score is a common linear transformation

$$Z = \frac{X - \mu}{\sigma}$$

- By "z scoring" a random variable, the new random variable (Z) will have mean 0 and variance 1.

The Relationship Between Two Random Variables

- Previously, we talked about the distribution, mean and variance for a single random variable.
- However, like the concepts of correlation and covariance *for data*, there are similar ideas for random variables.

The Joint Distribution Function

- When we deal with two random variables, X and Y, it is convenient to work with *joint probabilities*. We define the joint probability distribution to be

$$P_{X,Y}(x,y) = P(X = x \, and \, Y = y)$$

- As usual, we require that

$$P_{X,Y}(x,y) \geq 0 \text{ for any pairs x,y}$$

$$\sum_{\text{all x,y}} \mathrm{P}_{\mathrm{X,Y}}(x,y) = 1$$

Example: More Credit, More Purchases

- The accompanying table shows, for credit card holders with one to three cards, the joint probabilities for number of cards owned (X) and number of credit purchases made in a week (Y).

Number of Cards (X)	Number of Purchases in Week (Y)				
	0	1	2	3	4
1	0.08	0.13	0.09	0.06	0.03
2	0.03	0.08	0.08	0.09	0.07
3	0.01	0.03	0.06	0.08	0.08

Example: Stocks and Bonds

Treasury Bills (T) \ Stocks (S)	-10%	0	10%	20%	$P_T(t)$
6%	0	0	.10	.10	
8%	0	.10	.30	.20	
10%	.10	.10	0	0	
$P_S(s)$					

Marginal Distributions

- Suppose we are interested only in X, yet have to work with the joint distribution of X and Y. We can obtain the *marginal distribution* of X as follows.
- The marginal probabilities of X and Y are given by

$$P_X(x) = \sum_y P_{X,Y}(x,y) \quad and \quad P_Y(y) = \sum_x P_{X,Y}(x,y)$$

- As before, the term “marginal” merely describes how the distribution of X can be calculated from the joint distribution of X and another variable Y; row sums (or column sums) are calculated and placed “in the margin”

Example of Marginal Distributions

- Compute the marginal distributions

Stocks (S)

Treasury Bills (T)	-10%	0	10%	20%	$P_T(t)$
6%	0	0	.10	.10	
8%	0	.10	.30	.20	
10%	.10	.10	0	0	
$P_S(s)$					

Tbills	Prob
6%	
8%	
10%	

Stocks	Prob
-10%	
0	
10%	
20%	

Independence

- Two random variables X and Y are called **independent** if the events (X=x) and (Y=y) are independent. That is,
- The random variables X and Y are independent if for ***all* values** of *x* and *y*:

$$P_{X,Y}(X = x\, and\, Y = y) = P_X(x)P_Y(y)$$

Test for Independence

	Stocks (S)				
Treasury Bills (T)	-10%	0	10%	20%	$P_T(t)$
6%	0	0	.10	.10	
8%	0	.10	.30	.20	
10%	.10	.10	0	0	
$P_S(s)$					

Conditional Distributions

- Let X and Y be jointly distributed random variables. Then the **conditional distribution** of X given Y is given by

$$P_{X|Y}(X = x|Y = y) = \frac{P_{X,Y}(x, y)}{P_Y(y)}$$

Example: Stock Market Table

		Stocks (S)				
		-10%	0	10%	20%	$P_T(t)$
Treasury Bills (T)	6%	0	0	.10	.10	
	8%	0	.10	.30	.20	
	10%	.10	.10	0	0	
	$P_S(s)$					

Compare P(S) and P(S|T=10%)

Conditional Expectation

- One useful application of conditional distributions is in calculating conditional expectations. You will see a lot more of this when we get to regression analysis.

- The basic idea is that given a conditional distribution, we can also calculate a conditional expectation:

$$E(X|Y=y) = \sum_{all\,x\,values} xP(X=x|Y=y)$$

Combining Random Variables

- IF X and Y are independent, it is easy to combine random variables.
- That is, IF X and Y are independent,
- $E(X+Y)=E(X)+E(Y)$ [actually always true]
- $Var(X+Y)=Var(X)+Var(Y)$
- To understand how to calculate $E(X+Y)$ and $Var(X+Y)$ for all scenarios, we need to first introduce the concepts of covariance and correlation for random variables.

Covariance and Correlation

- The variance of a random variable is a measure of its variability, and the covariance of two random variables is a measure of their joint variability.
- The covariance is a measure of the *linear association* of two random variables. Its sign reflects the direction of the association; if the variables tend to move in the same direction the covariance is positive. If the variables tend to move in opposite directions the covariance is negative.

Covariance is a pain to calculate

- The covariance is a bit of a pain to calculate.

$$\boxed{\sigma_{XY} = \sum_{i=1}^{N} [x_i - E(X)][(y_i - E(Y)]\, P(x_i, y_i)}$$

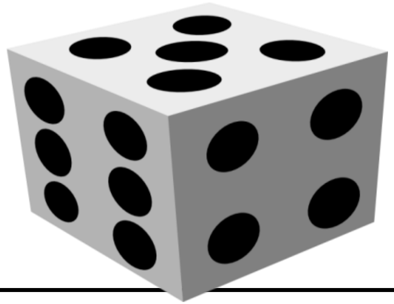
- Three interesting facts: (1) Cov(X,X)=Var(X),

  (2) **if X and Y are independent**, Cov(X,Y)=0,

  (3) Cov(X,Y)=E(XY)-E(X)E(Y) (<u>alternative formula</u>)

Example

- Consider a 6 sided die and let X be the number on the top of the die and Y be the number on the bottom.
- Does anyone know the relationship between X and Y?



Calculate the Covariance

- We will use the formula Cov(X,Y)=E(XY)-E(X)E(Y)
- For a die E(X)=E(Y)=3.5
- We need to find E(XY)

Probability	X	Y	XY	Prob × XY
1/6	1	6	6	6/6 = 1
1/6	2	5	10	10/6 = 5/3
1/6	3	4	12	12/6 = 2
1/6	4	3	12	12/6 = 2
1/6	5	2	10	10/6 = 5/3
1/6	6	1	6	6/6 = 1
				$E(XY) = \text{sum} = 9\frac{1}{3} = 9.333$

- So Cov(X,Y)=9.33-(3.5)(3.5) = -2.91
- The covariance is negative because larger values of X are associated with smaller values of Y.

Example: Stocks and Bonds

		Stocks (S)				
		-10%	0	10%	20%	$P_T(t)$
Treasury Bills (T)	6%	0	0	.10	.10	
	8%	0	.10	.30	.20	
	10%	.10	.10	0	0	
	$P_S(s)$					

Example using the longer formula for covariance

$$E(S) = -10(.1) + 0(.2) + 10(.4) + 20(.3) = 9$$

$$E(T) = 6(.2) + 8(.6) + 10(.2) = 8$$

$$\begin{aligned} Cov(S,T) = & (-10-9)(10-8)(.1) \\ & +(0-9)(8-8)(.1) + (0-9)(10-8)(.1) \\ & +(10-9)(6-8)(.1) + (10-9)(8-8)(.3) \\ & +(20-9)(6-8)(.1) + (20-9)(8-8)(.2) \\ & = -9.1 \end{aligned}$$

The Covariance Matrix

- Sometimes the covariance between random variables is presented in a table, or matrix of the following form:

	X1	X2	X3
X1	Var(X1)	Cov(X1,X2)	Cov(X1,X3)
X2	Cov(X2,X1)	Var(X2)	Cov(X2,X3)
X3	Cov(X3,X1)	Cov(X3,X2)	Var(X3)

this is called a covariance matrix

Correlation: Covariance Rescaled

- Covariance can indicate whether X and Y have a positive, negative, or zero relation. Yet, it turns out not to be a good measure of association since it depends on the units of measurement.
- To eliminate this difficulty, we define the correlation:

$$\rho = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y}$$

- This is a dimensionless measure of association.
- The correlation is always between -1 and 1, with 1 indicating a perfect positive linear relationship, -1 a perfect negative linear relationship and 0 no linear relationship between X and Y.

Dice Example Again

- X top of dice, Y= bottom of dice
- E(X)=3.5 and Var(X) = 2.91 (same for Y)
- We found earlier that Cov(X,Y) = -2.91
- Then the correlation is

$$\rho = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y} = \frac{-2.91}{\sqrt{2.91}\sqrt{2.91}} = -1$$

- This makes sense since X= 7-Y (perfect negative relationship)

Combinations of Random Variables

- If X and Y are independent

$$E(X+Y)=E(X)+E(Y)$$
$$Var(X+Y)=Var(X)+Var(Y)$$

- If X and Y are not independent

$$E(X+Y)=E(X)+E(Y)$$
$$Var(X+Y)=Var(X)+Var(Y)+2Cov(X,Y)$$

- The most general case

$$E((a+bX)+(c+dY))=a+bE(X)+c+dE(Y)$$
$$Var((a+bX)+(c+dY))=b^2Var(X)+d^2Var(Y)+2bdCov(X,Y)$$

Dropping Coins

- Suppose a game involves dropping 3 coins on the table—a nickel, a dime, and a quarter. Each coin that lands "heads up" you are allowed to keep, so that the possible reward R ranges from 0 to 40c.
- Find the mean and variance of R.
- This can be done by brute force by first finding the distribution of R, then the mean and variance.
- But using the idea of sums of random variables, there is an easier approach.

Breaking down the problem

- Define the following random variables
  - X1 = the nickel's contribution to the reward
  - X2 = the dime's contribution to the reward
  - X3 = the quarter's contribution to the reward
- Then R = X1 + X2 + X3
- E(X1) = .5(.05)+.5(0) [why??]
- E(X2) = .5(.1)+.5(0)
- E(X3) = .5(.25)+.5(0)
- E(R) = E(X1)+E(X2)+E(X3) = 0.20
- How do we find the variance??

The Hat Check Problem

- There is a dinner party where n men check their hats.
- The hats are mixed up during dinner, so that afterward each man receives a random hat. In particular, each man gets his own hat with probability 1/n.
- What is the expected number of men who get their own hat?

The Hat Check Problem

- Letting G be the number of men that get their own hat, we want to find E(G).
- But all we know is that the probability a man gets his own hat back is 1/n.
- We don't know enough about the distribution of G to calculate its expectation directly.
- But linearity of expectation makes the problem really easy.

The Hat Check Problem

- Define the random variables

$$G_i = \begin{cases} 1 & \text{i'th person gets their hat back} \\ 0 & \text{otherwise} \end{cases}$$

- Note that

$$E(G_i) = 1P(G_i = 1) + 0P(G_i = 0) = 1/n$$

- Also, these indicators are not mutually independent!

The Hat Check Problem

- The number of men that get their own hat is the sum of these indicators:

$$G = G_1 + G_2 + \cdots + G_n$$

- Then E(G) is easily found as

$$E(G) = E(G_1 + G_2 + \cdots + G_n)$$
$$= E(G_1) + E(G_1) + \cdots + E(G_n)$$
$$= n(1/n) = 1$$

So even though we don't know much about how hats are scrambled, we've figured out that on average, just one man gets his own hat back!

Example: Stock Prices

- Suppose your portfolio consists of \$1000 under your mattress, 5 shares of Stock X and 10 shares of stock Y.
- The value (wealth) of your portfolio is then

$$W=1000 + 5X + 10Y$$

Joint probability table

- Suppose next months prices of the two stocks are modeled as the following joint probability table

			Stock Y Price		
		40	**50**	**60**	**70**
Stock	**45**	0.24	0.003333	0.003333	0.003333
X	**50**	0.003333	0.24	0.003333	0.003333
Price	**55**	0.003333	0.003333	0.24	0.003333
	60	0.003333	0.003333	0.003333	0.24

Find all the expected value and variances

- It may be shown (do this yourself-good practice!)
- $E(X) = 53$ and $E(Y) = 55$
- $Var(X) = 31.3$ and $Var(Y) = 125$
- $COV(X,Y) = 59.17$

Find E(W) and Var(W)

- Using

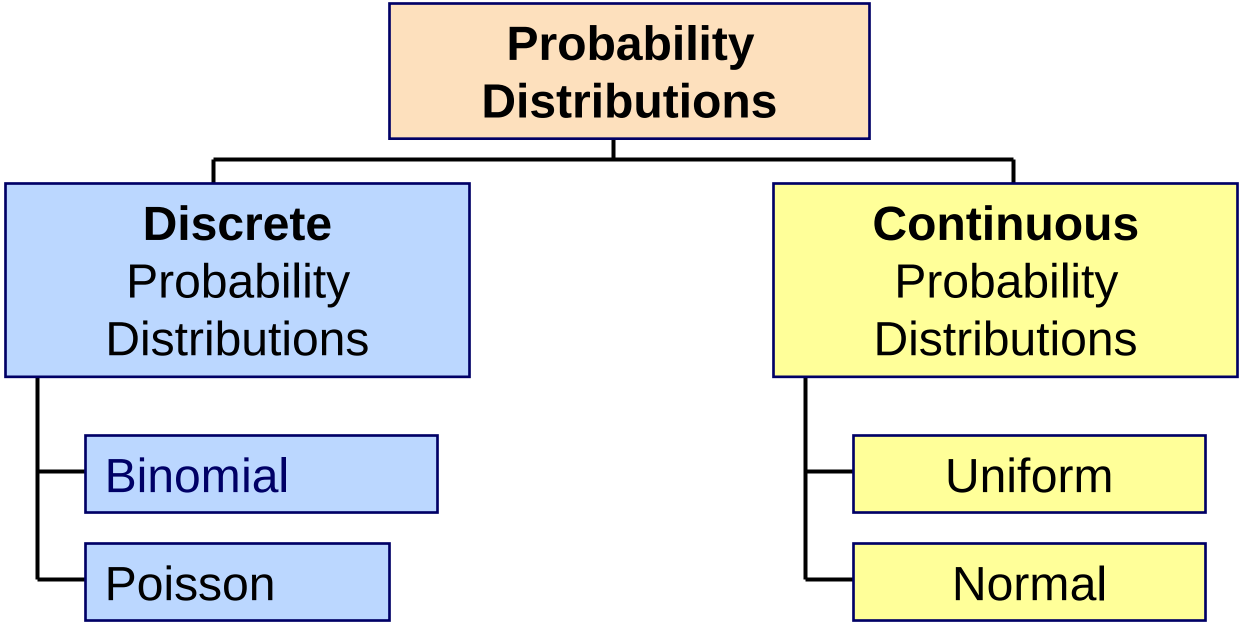
$$E((a+bX)+(c+dY)) = a + bE(X) + c + dE(Y)$$

$$Var((a+bX)+(c+dY)) = b^2Var(X) + d^2Var(Y) + 2bdCov(X,Y)$$

$$\begin{aligned} E(W) &= 1000 + 5E(X) + 10E(Y) \\ &= 1000 + 5(53) + 10(55) \\ &= 1815 \end{aligned}$$

$$\begin{aligned} Var(W) &= 5^2Var(X) + 10^2Var(Y) + 2(5)(10)Cov(X,Y) \\ &= 25(31.3) + 100(125) + 100(59.17) \\ &= 19199.5 \end{aligned}$$

Probability Distributions



Probability distributions and random variables go hand in hand. We talk for example about the binomial probability distribution or a binomial random variable-same thing.

The Binomial Distribution

- Many business experiments can by characterized by a binomial distribution (random variable).
- The key ideas are as follows:
  - We are doing something *n* times.
  - Each instance of a trial has only two outcomes, “success” or “failure”.
  - The trials of the experiment are independent of each other.
    - The probability of a success, *p*, remains constant from trial to trial
    - We are interested in the **total number of successes** (out of *n* trials)

Binomial random variable is for modeling counts.

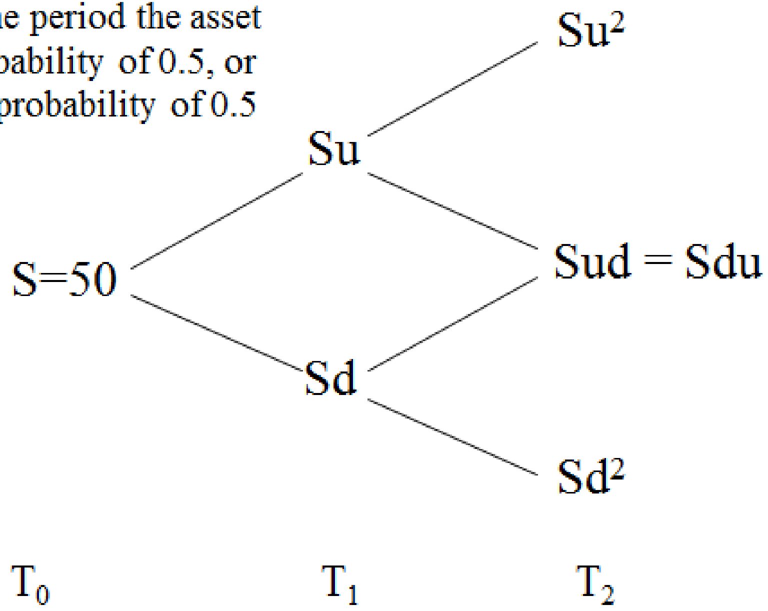
Binomial Tree of Asset Prices

- The most common application of the binomial distribution in finance is ‘security price change’
- It is assumed that over the next small interval of time security price will wither rise (‘a success’) or fall (‘a failure’) by a given amount
- The binomial distribution is an assumption in some option pricing models

Binomial Tree

A binomial tree of asset prices

In each of the time period the asset may rise with probability of 0.5, or it may fall with a probability of 0.5



Binomial Distribution Formula

$$P(X=x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

x = number of 'successes' in sample,
(x = 0, 1, 2, ..., n)

p = probability of "success" per trial

q = probability of "failure" = (1 – p)

n = number of trials (sample size)

Example

- In the game of blackjack as played in Las Vegas casinos, the dealer has the advantage. Since most players don't play well, the probability that the average player wins a hand is about 45%. Find the probability that an average player wins ten or more times in 25 hands.

```
 1-pbinom(9,25,.45)
[1] 0.7576288

pbinom(x,n,p) cumulative probablility F(x)
dbinom(x,n,p) P(x)
```

Discrete versus Continuous RVs

The major distinction between a continuous and discrete random variable is the numerical events of interest.

We can list all possible values of a discrete random variable, and it is meaningful to consider the probability that a particular individual value will be assumed.

We cannot list all the values of a continuous random variable-because there is always another possible value between any two of its values.

Hence the only meaningful events for a continuous random variable are *intervals.*

The probability that a continuous random variable X will assume any particular value is 0.

P(X=x) = 0 for any x if X is a continuous random variable

Hence, for a continuous random variable X, it is only meaningful to talk about the probability that the value assumed by X will fall within some *interval of values*.

$$P(a \le X \le b)$$

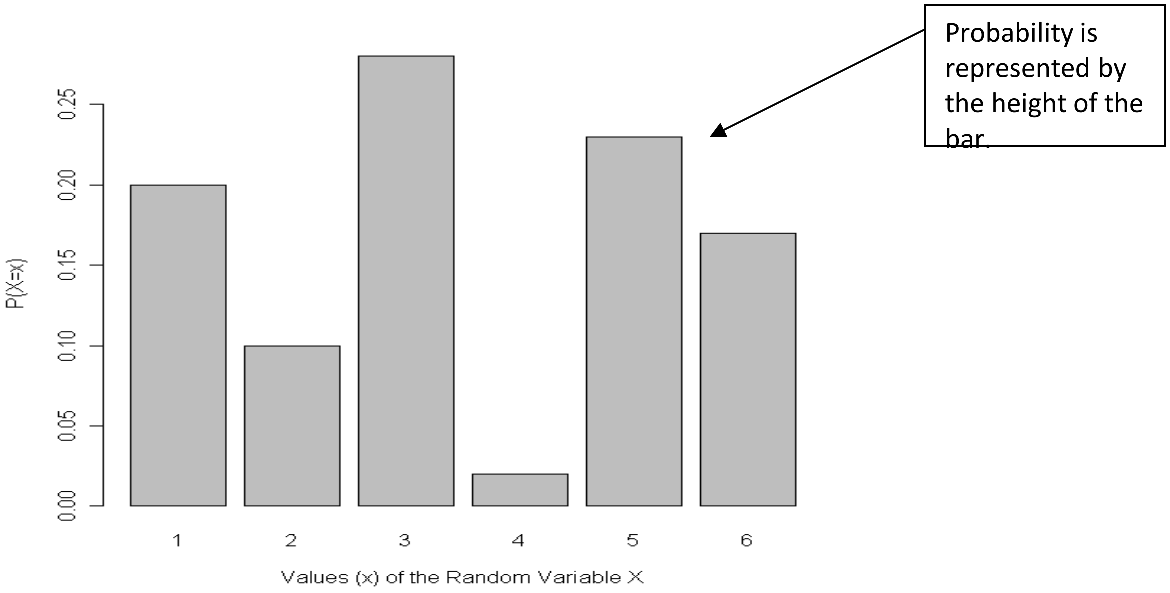
Note that

$$P(a \le X \le b) = P(a < X < b) = P(a \le X < b)[why?]$$

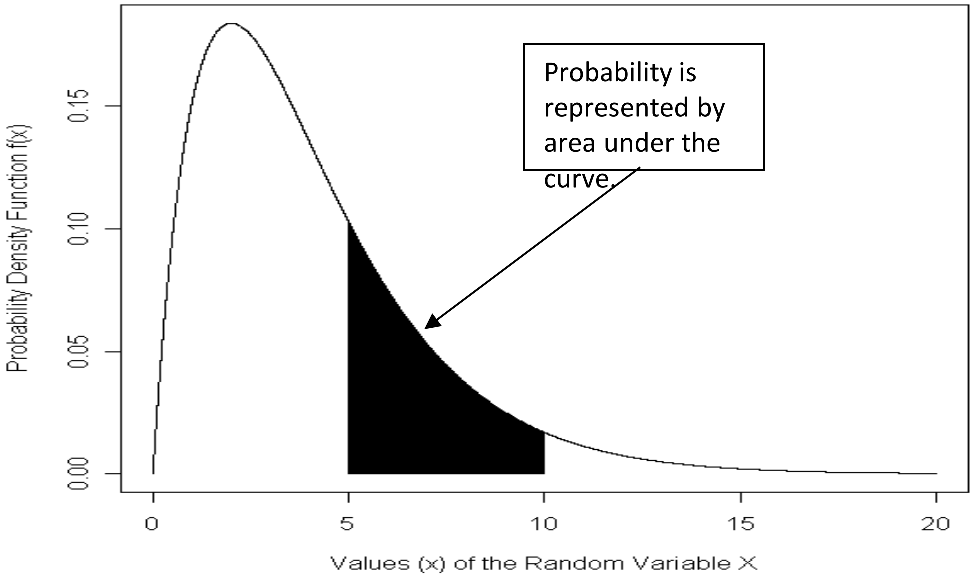
Density Curves

- Graphs are usually the easiest way to describe a continuous random variable.
- The **probability density function** ***f(x)*** is a curve that describes the probability associated with the range of values that a continuous random variable can assume.
- The <u>area under the curve</u> of the density function <u>is the probability</u> that any range of values will occur. The total area under the curve has to be one.

Graph of a Discrete Distribution



Graph of a Continuous Distribution



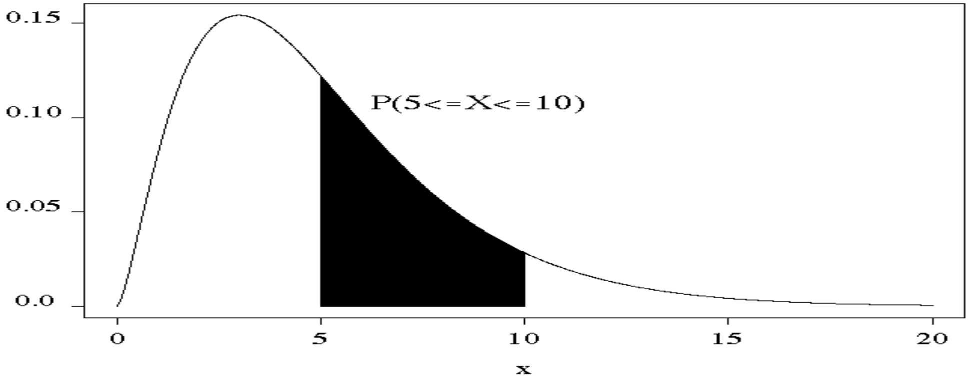
The density curve f(x) is such that f(x)>=0 for all x and the area under the curve=1

The Density Curve f(x)

- Notice that the vertical scale is labeled *f(x)*, not *P(X=x).*
- This difference in notation emphasizes the fact that the height in a continuous distribution is not a direct measure of probability.
- To **produce probabilities** in the continuous case, we'll need to **measure areas under the curve**. We usually will need a computer or special tables to find the desired area under the curve.
- What is often handy though is to visually inspect the curve to understand what values are likely to occur.
- If we examine the last figure, its clear that values larger than 15 are unlikely (very little area under the curve) and negative values of the random variable are not possible at all (why?).

More on the Density Curve

- As previously mentioned, the probability that X will take any specific value is zero. Given a pdf f(x), the area under the graph of f(x) between the two values a and b is the probability that X will take a value between a and b. In calculus notation, we have that



$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

=area under curve between a and b

CDF: Cumulative Distribution Function

- The cumulative distribution function (cdf) is useful because it gives us the probability of being less than some value:

$$F_X(X \leq x) = P(X \leq x) = \int_{-\infty}^{x} f(x)dx$$

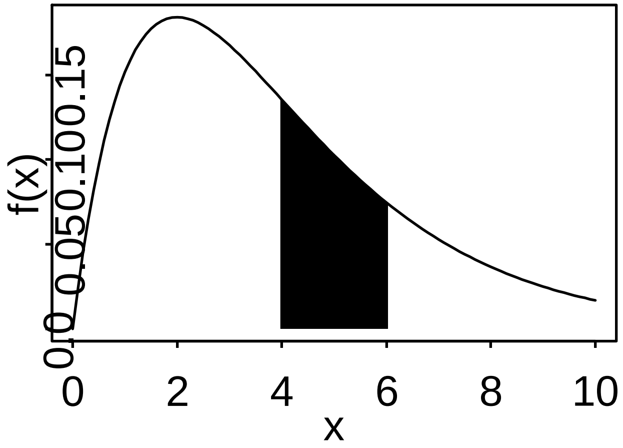
- We will never calculate this directly; we usually obtain these values from the computer or a table.

Probability of an Interval

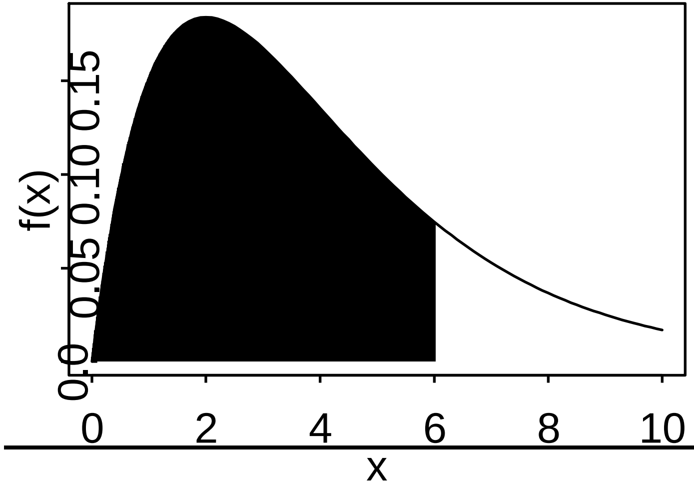
- Using the CDF, we can calculate the probability of any interval:

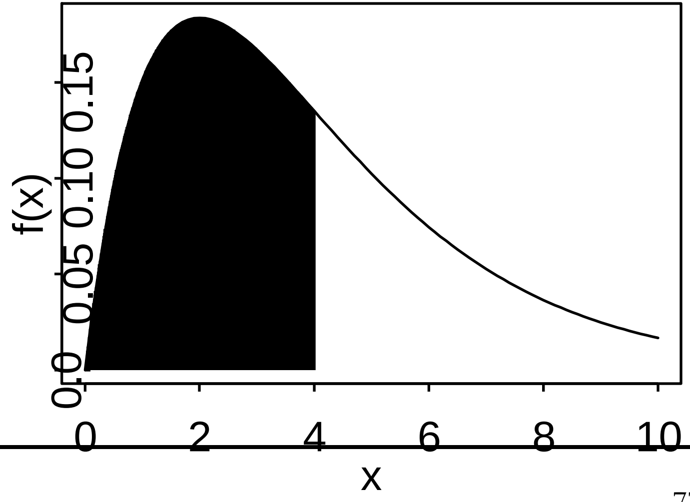
$$P(a \leq X \leq b) = F_X(b) - F_X(a)$$

To find this shaded area, subtract the two shaded areas below.



=





Mean of a continuous RV

- The expected value of a continuous random variable *X* is defined to be

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

- This is just the continuous version of

$$\mu_X = E(X) = \sum_{all\ x_i} x_i P_X(x_i)$$

This formula is given for completeness-we don't assume calculus knowledge.

Variance of a continuous RV

- The variance of a continuous random variable *X* is defined to be

$$\sigma^2 = Var(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx$$

- This is just the continuous version of

$$\sigma_X^2 = \sum_{all\ x_i} (x_i - \mu)^2 P(X = x_i)$$

This formula is given for completeness-we don't assume calculus knowledge.

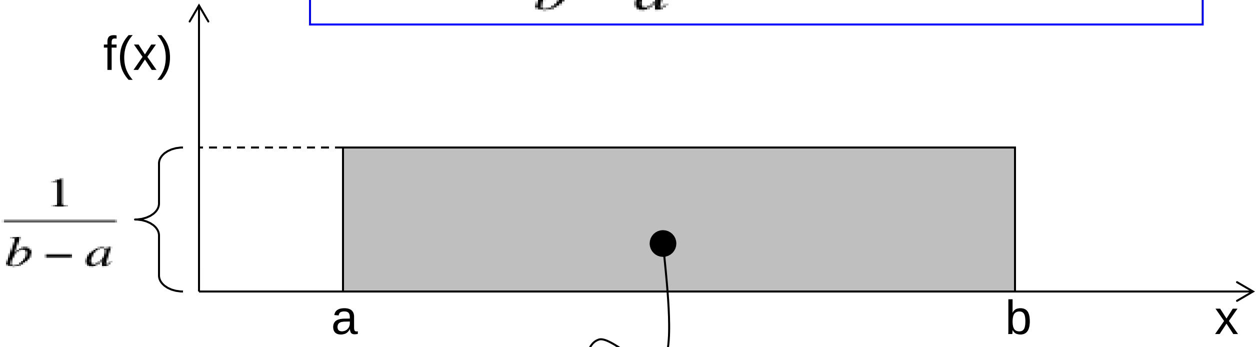
The Uniform Distribution

- One of the simplest examples of a continuous probability distribution is the uniform distribution.
- We'll use it here to illustrate typical continuous probability distribution characteristics.

The Uniform Distribution

- Consider the ***uniform probability distribution***
- It is described by the function:

$$\boxed{f(x) = \frac{1}{b-a}, \quad where \quad a \le x \le b}$$



area = width x height = (b – a) x $\frac{1}{b-a}$ = 1

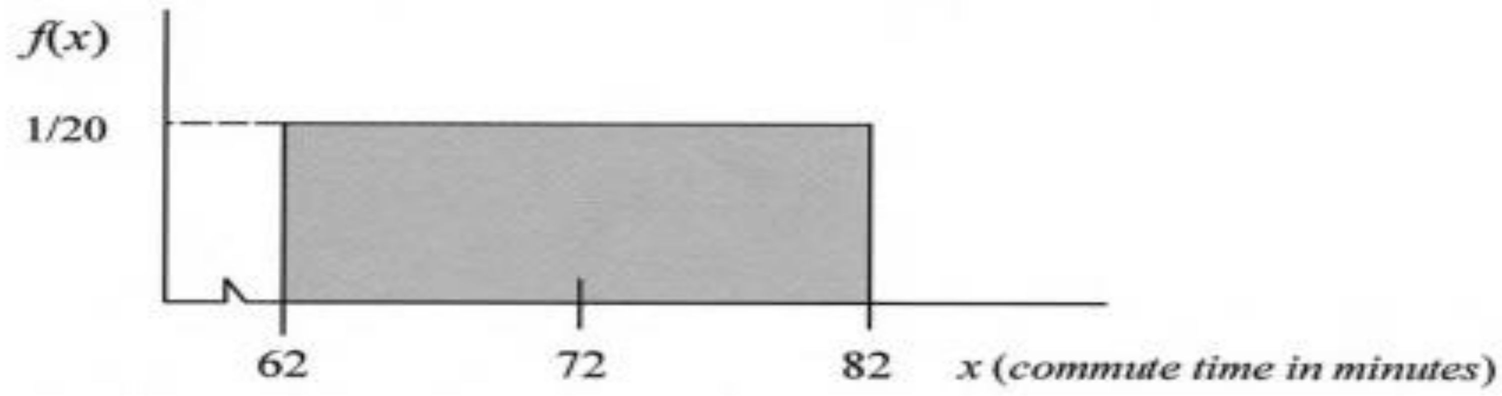
Example

- Lyla's daily commute time is a random variable— call it X— with a uniform probability distribution defined by the probability density function

$$f(x) = \begin{cases} \dfrac{1}{20} & \text{for x values between 62 and 82 minutes} \\ 0 & \text{everywhere else} \end{cases}$$

A Picture

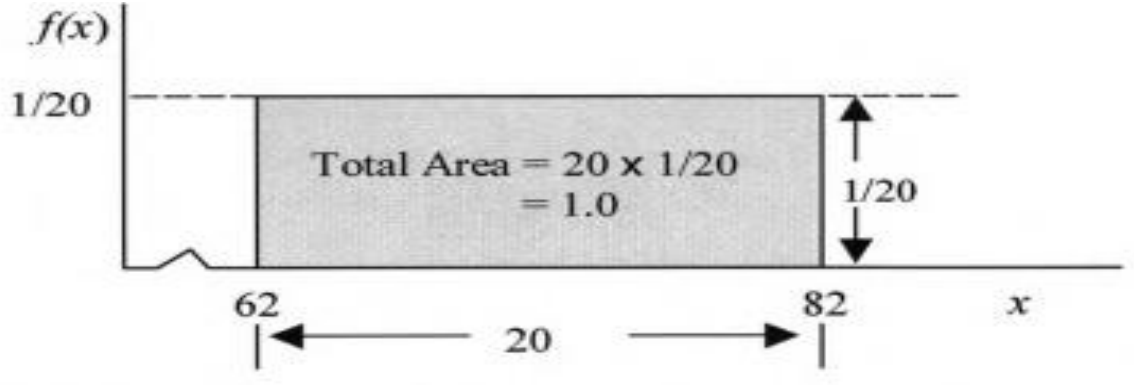
- The uniform probability distribution here shows that Lyla's commute time can be anywhere between 62 and 82 minutes and is just as likely to be in the interval 63 to 64 minutes as in the interval 64 to 65 minutes.



- It's this "equal probability for equal intervals" characteristic that identifies any uniform probability distribution.

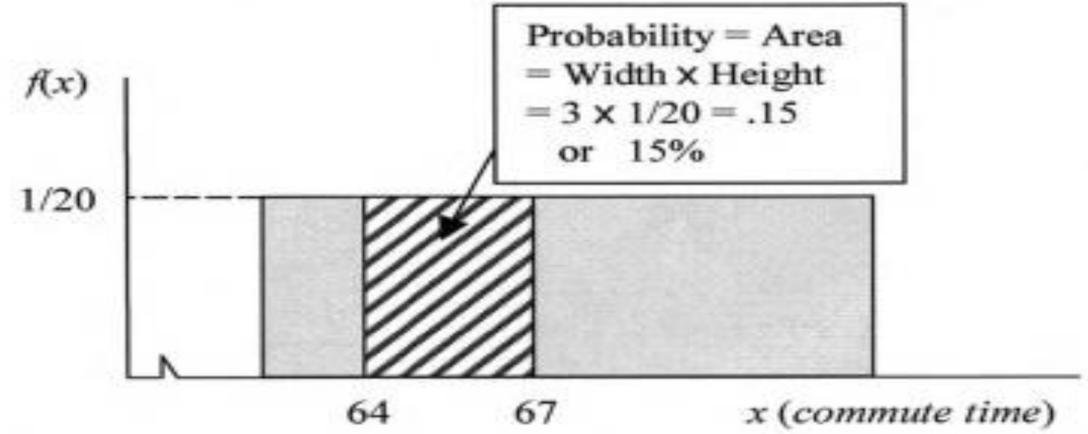
Total area under the curve

- The total area under the curve is (and has to be) 1.
- This is equivalent to the fact that the sum of probabilities for a discrete distribution has to be one.



Finding Probabilities

- We want to find the probability that Lyla's commute is between 64 and 67 minutes.
- This is simply the area under the curve between those 2 points



General Characteristics

- In general the uniform density is given by

$$f(x) = \begin{cases} \dfrac{1}{(b-a)} & \text{for x values between a and b minutes} \\ 0 & \text{everywhere else} \end{cases}$$

- The mean is then

$$E(X) = \frac{(a+b)}{2}$$

- The variance is

$$Var(X) = \frac{(b-a)^2}{12}$$

The Normal Distribution

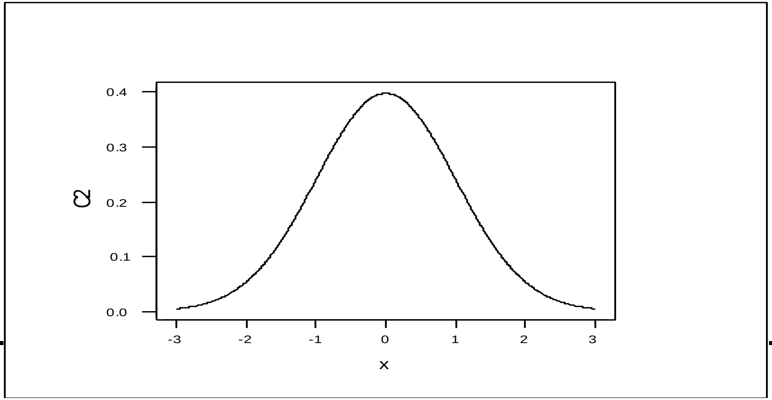
The normal distribution is the most fundamental continuous distribution used in statistics so I guess we better learn it. Many methods that are widely used in economics, finance, and marketing are based on an assumption of a normal distribution.

Normal Dist.

Gaussian Dist.

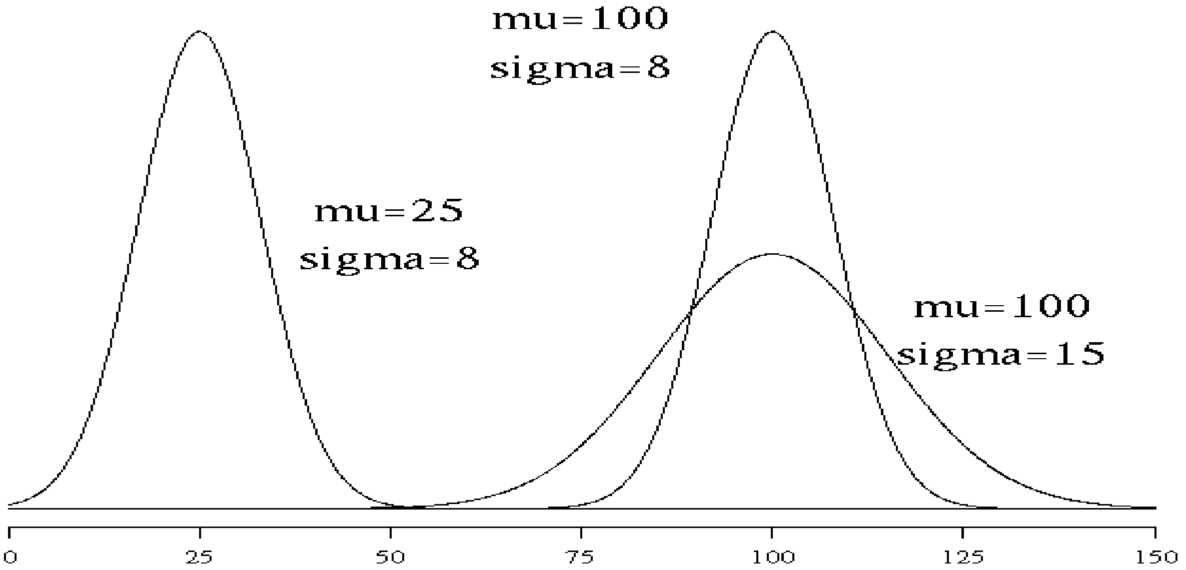
Bell-Shaped Curve

equivalent names



What controls the shape of the curve ?

The normal distribution is governed by the **two parameters :** μ (the *mean*) and σ (the *standard devia*

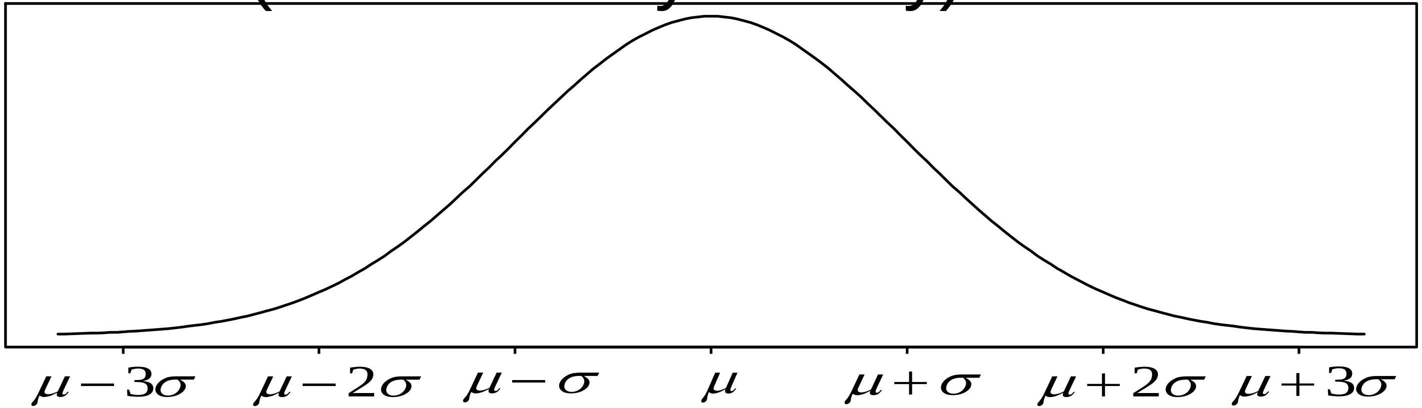


Notation for the normal curve

We use the notation

$$X \sim N(\mu, \sigma^2).$$

Here is a general picture of a normal distribution (notice the symmetry).



Using the Computer

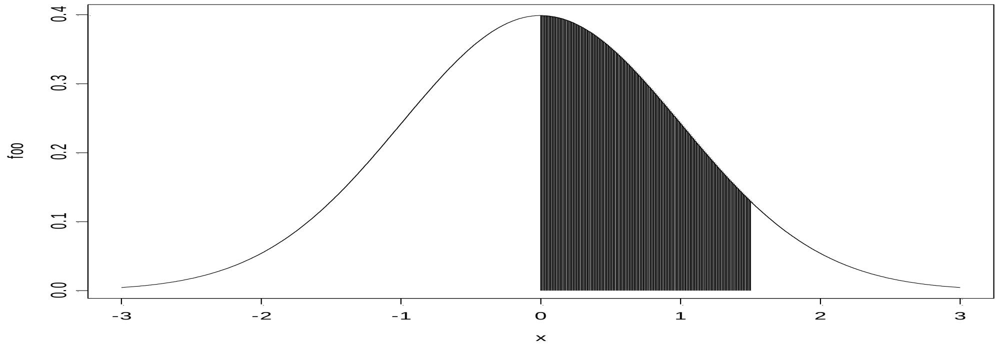
- R
  - P(X<x) = pnorm `(x,m,s)`
  - `f(x) = dnorm(x,m,s)`

Example

- Daily stock market returns on the NY Stock Exchange are approximately normally distributed, with a mean of 0.0317% and a standard deviation of 1.046%.
- What percentage of the time are daily returns positive?
- What percentage of the time do daily returns return between 1% and 2%?

Finding normal probabilities

- $P(a \leq X \leq b)$ = area under the curve between *a* and *b*.



We use the computer or tables to do this

Example

- We are told X is normal m= 0.0317 and s=1.046
- To find P(X>0)

```
> 1-pnorm(0,.0317,1.046)
[1] 0.5120885
```

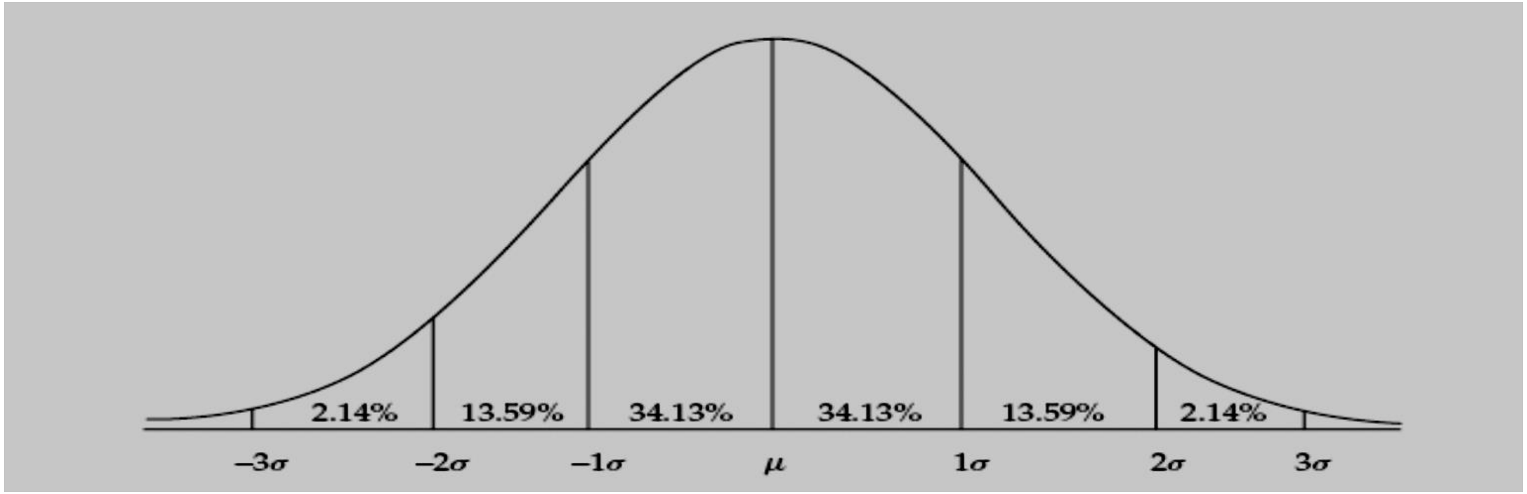
Example

- What percentage of the time do daily returns return between 1% and 2%?

```
pnorm(2,.0317,1.046)-pnorm(1,.0317,1.046)
[1] 0.1473609
```

Normal Coverage Rule

$$\begin{aligned}
&For\ X \sim N(\mu, \sigma^2),\\
&P(\mu - \sigma \le X \le \mu + \sigma) = 0.68\\
&P(\mu - 2\sigma \le X \le \mu + 2\sigma) = 0.95\\
&P(\mu - 3\sigma \le X \le \mu + 3\sigma) = 0.997
\end{aligned}$$



Returns may not be Normal

- It is nice to assume returns are normally distributed but there is an issue
- Simple returns are defined to be

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

- As long as the price>0, $R_t > -1$ [the most you can lose is all your money]

Returns may not be Normal

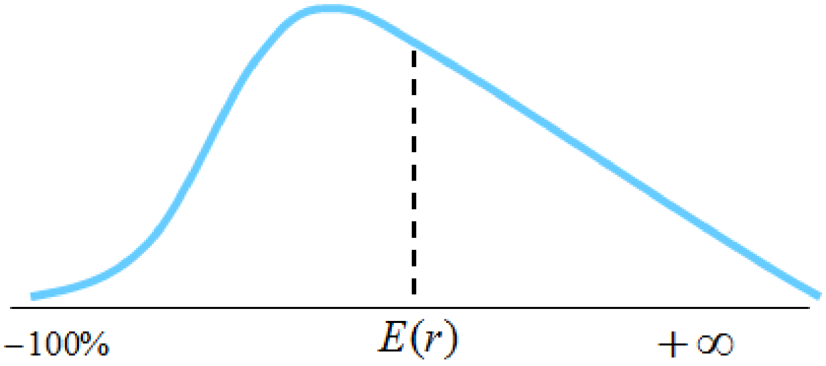
- Suppose returns are normal with mean 0.05 and std dev .5
- Lets calculate $P(R_t < -1)$

```
pnorm(-1,.05,.5)
[1] 0.01786442
```

- This implies that there is a 1.8% chance that the asset price will be negative. This is why the normal distribution may not be appropriate for simple returns.

The Log Normal Distribution

- The Log Normal distribution is sometimes used to compensate for this-more details later on.



Skewness-measure of symmetry

- Skewness of a random variable is

$$\text{Skew}(X) = E\left[\left(\frac{X-\mu_X}{\sigma_X}\right)^3\right]$$

- If X has a symmetric distribution about μ then Skew=0
- Skew > 0 ⇒ pdf has long right tail, and median < mean
- Skew < 0 ⇒ pdf has long left tail, and median > mean

Skewness Example

- Using the discrete distribution for the return on Microsoft stock shown earlier, we have

$$
\begin{aligned}
\text{skew}(X) &= [(-0.3 - 0.1)^3 \cdot (0.05) + (0.0 - 0.1)^3 \cdot (0.20) \\
&+ (0.1 - 0.1)^3 \cdot (0.5) + (0.2 - 0.1)^3 \cdot (0.2) \\
&+ (0.5 - 0.1)^3 \cdot (0.05)]/(0.141)^3 \\
&= 0.0
\end{aligned}
$$

Kurtosis-measure heavy tails

- Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution.
- That is, data sets with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails.
- Data sets with low kurtosis tend to have a flat top near the mean rather than a sharp peak.
- Normal data has kurtosis=3 (or 0 depending on routine).

Kurtosis-measure of tail thickness

- Defined to be

$$\mathrm{Kurt}(X) = E\left[\left(\frac{X-\mu_X}{\sigma_X}\right)^4\right]$$

- For our MSFT example

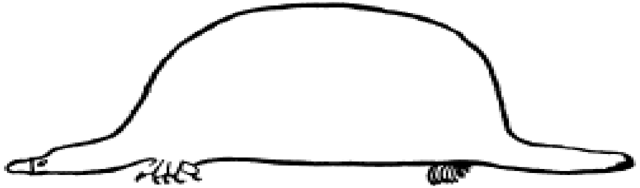
$$\begin{aligned}\mathrm{Kurt}(X) &= [(-0.3-0.1)^4\cdot(0.05)+(0.0-0.1)^4\cdot(0.20)\\ &+(0.1-0.1)^4\cdot(0.5)+(0.2-0.1)^4\cdot(0.2)\\ &+(0.5-0.1)^4\cdot(0.05)]/(0.141)^4\\ &= 6.5\end{aligned}$$

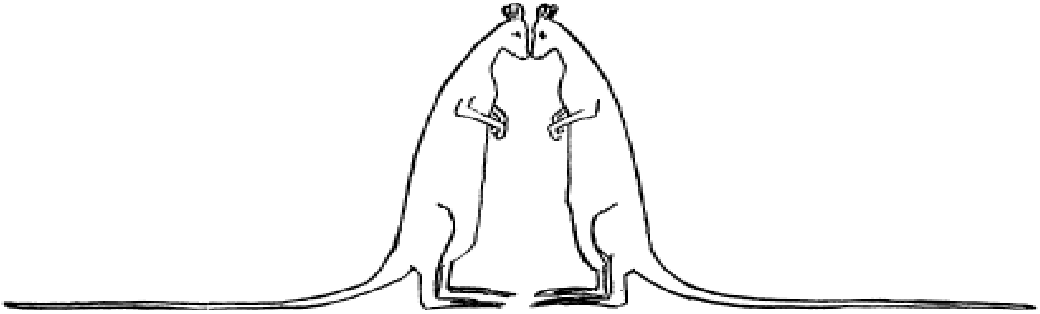
Kurtosis compared to 3

- A normal distribution has a kurtosis of 3
- Kurtosis >3 $\Rightarrow$ X has fatter tails than normal distribution
- kurtosis <3 $\Rightarrow$ X has thinner tails than normal distribution

No Comment

* In case any of my readers may be unfamiliar with the term "kurtosis" we may define mesokurtic as "having β_2 equal to 3," while platykurtic curves have $\beta_2 < 3$ and leptokurtic > 3. The important property which follows from this is that platykurtic curves have shorter "tails" than the





normal curve of error and leptokurtic longer "tails." I myself bear in mind the meaning of the words by the above *memoria technica*, where the first figure represents platypus, and the second kangaroos, noted for "lepping," though, perhaps, with equal reason they should be hares!

Introduction to Value at Risk

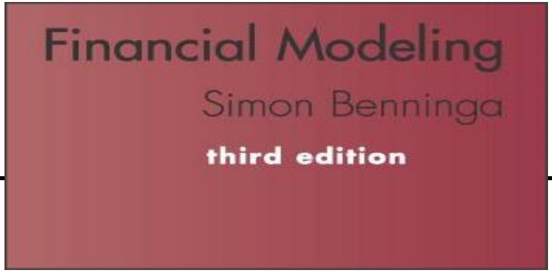
- Consider a Wo = $10000 investment in MSFT for one month.
- Assume the return is normally distributed with mean 0.05 and stdev 0.1

$$R \sim N(0.05, (0.1)^2)$$

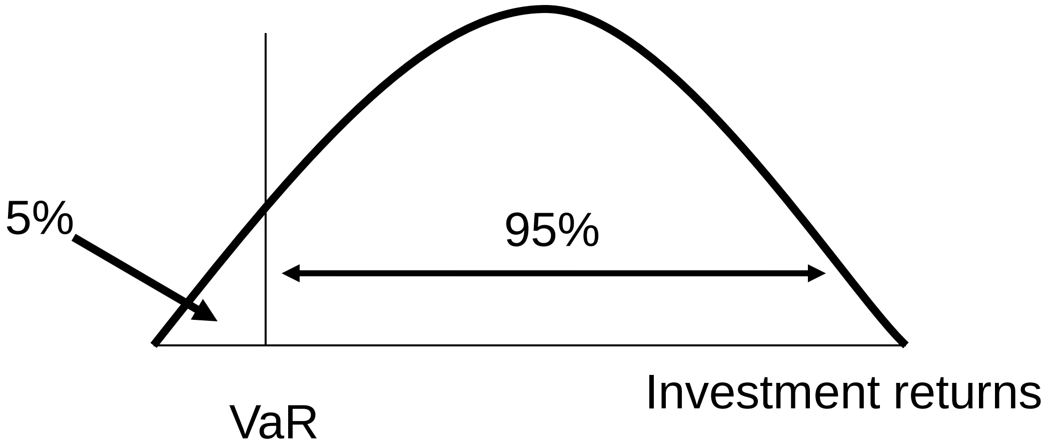
- Goal-calculate how much we can lose with specified probability α

In the Jargon of VaR

- In the jargon of VaR, suppose that a portfolio manager has a daily VaR equal to $1 million at 1 percent. This statement means that there is only one chance in 100 that a daily loss bigger than $1 million occurs under normal market conditions.

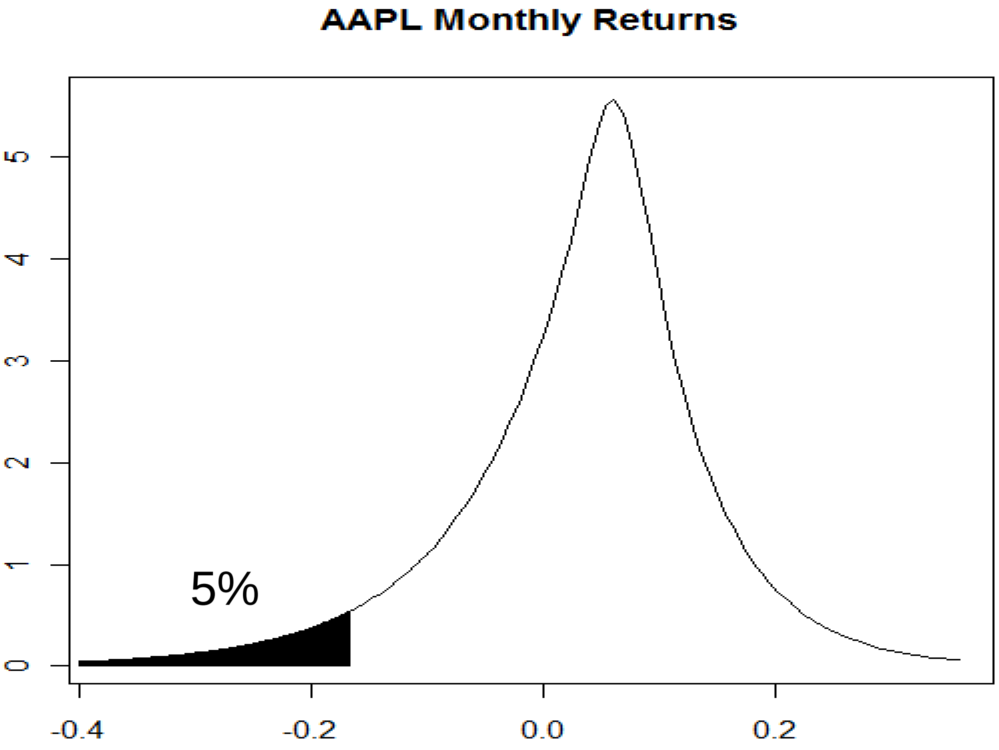


Choice of confidence level – 95%



Normal market conditions – the returns that account for 95% of the distribution of possible outcomes.
Abnormal market conditions – the returns that account for the other 5% of the possible outcomes.

Consider AAPL



Value at Risk

- What is the monthly value-at-risk (VaR) on the $10,000 investment with 5% probability?

$$R \sim N(0.05, (0.1)^2)$$

```
> qnorm(.05,.05,.1)
 [1] -0.1144854
```

- There is a 5% chance of a loss of 11.44% or worse.
- The loss in investment value is $10000(-11.44) = -$1144
- Because VAR represents a loss it is usually reported as a positive number.