Statistics and probability in equity analysis I

FN 451

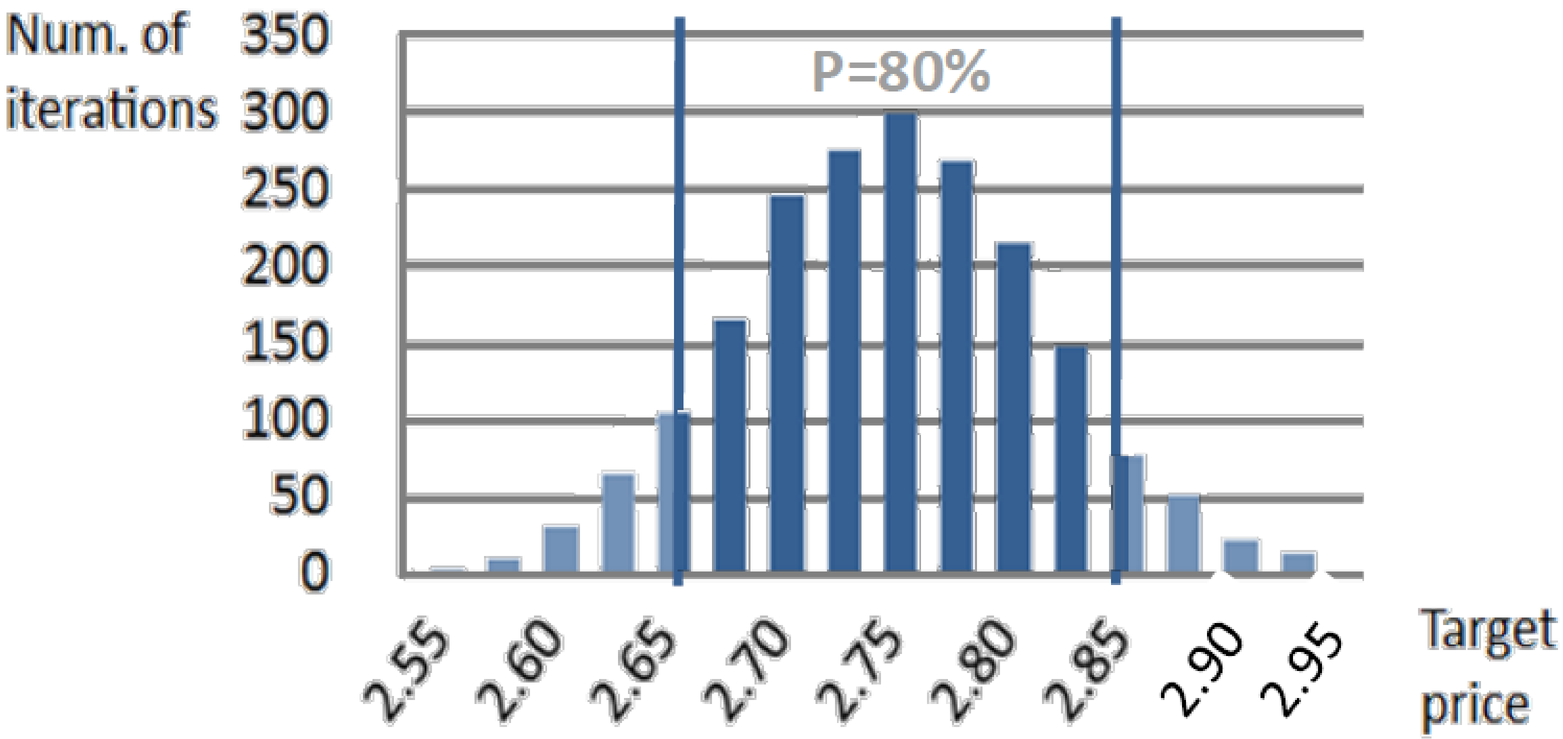
Class 3

THAMMASAT
BUSINESS SCHOOL

Outline

- Mathematics of risk and return
- Types of returns : Evaluation of investment performance
  - Geometric vs Arithmetic, Holding period returns
- Quantifying risk and uncertainty
  - Probability Theory
  - Normal distribution vs standard normal distributions
- Applications: Valuation simulations, technical analysis, portfolio risk management, portfolio allocation, performance attributions

Price simulations



THAMMASAT
BUSINESS SCHOOL

Questions that require knowledge or probability

- What is the likelihood of a stop-loss order being executed for my client's portfolio?
- What are the chances that the price of CPN will go up before mid-day trading break?

Why do we care?

- Process
- An exercise is logic and critical thinking
- Goal
- To come up with objective assessment on the firm's/market future performance.

RETURN AND RISK

Simple net return vs gross return

- Simple net return on asset dates t-1 and t is defined as,

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

- Simple gross return is defined as 1 + Rt
- Gross return over k periods from dates *t – k* to *t* as

$$1 + R_t(3) = (1 + R_1)\cdot(1 + R_2)\cdot(1 + R_3) = \frac{P_t}{P_{t-1}}\cdot\frac{P_{t-1}}{P_{t-2}}\cdot\frac{P_{t-2}}{P_{t-3}}$$

$$1 + R_t(k) = (1 + R_1)\cdot(1 + R_2)\cdots(1 + R_3) = \frac{P_t}{P_{t-1}}\cdot\frac{P_{t-1}}{P_{t-2}}\cdots\frac{P_{t-k+1}}{P_{t-k}} = \frac{P_t}{P_{t-k}}$$

Ways to annualize returns

- GM

$$Annualized\left[R_t(k)\right]=\left[\prod_{t=1}^{k}(1+R_t)\right]^{1/k}-1$$

- AM

$$Annualized\left[R_t(k)\right]\cong\frac{1}{k}\sum_{t=1}^{k}R_t$$

Will these two computation methods give close estimates?
Suppose k represents 1 year. You have a time series of return from year k = 1,…, to K.
What if the units of time change?

AM vs GM

Year	Payoff	Return
0	-100	
1	50	-50%
2	100	100%

AM = [(-50%) + 100%]/2 = 25%

GM = $[(1-0.5)(1+1)]^{1/2} - 1 = (0.5x2)^{1/2} - 1 = 0$

THAMMASAT
BUSINESS SCHOOL

AM vs GM: Motorola

	Revenues	% Change	EBITDA	% Change	EBIT	% Change
1994	$ 22,245		$ 4,151		$ 2,604	
1995	$ 27,037	21.54%	$ 4,850	16.84%	$ 2,931	12.56%
1996	$ 27,973	3.46%	$ 4,268	-12.00%	$ 1,960	-33.13%
1997	$ 29,794	6.51%	$ 4,276	0.19%	$ 1,947	-0.66%
1998	$ 29,398	-1.33%	$ 3,019	-29.40%	$ 822	-57.78%
1999	$ 30,931	5.21%	$ 5,398	78.80%	$ 3,216	291.24%
Arithmetic Average		7.08%		10.89%		42.45%
Geometric Average		6.82%		5.39%		4.31%
Standard deviation		8.61%		41.56%		141.78%

Source: Damodaran: Investment valuation, Tools Techniques for valuing any asset

From Gross Return to Holding-Period Returns: Geometric

- Recall gross return is

$$1+R_t(3)=(1+R_1)\cdot(1+R_2)\cdot(1+R_3)=\frac{P_t}{P_{t-1}}\cdot\frac{P_{t-1}}{P_{t-2}}\cdot\frac{P_{t-2}}{P_{t-3}}$$

$$1+R_t(k)=(1+R_1)\cdot(1+R_2)\cdots(1+R_3)=\frac{P_t}{P_{t-1}}\cdot\frac{P_{t-1}}{P_{t-2}}\cdots\frac{P_{t-k+1}}{P_{t-k}}=\frac{P_t}{P_{t-k}}$$

- The holding period return is the return that an investor would get when holding an investment over a period of *n* years, when the return during year *i* is given as R_i:

$$HPR=(1+R_1)\cdot(1+R_2)\cdots(1+R_n)-1$$

THAMMASAT BUSINESS SCHOOL

Holding Period Return: Example

- Suppose your investment provides the following returns over a four-year period:

Year	*Return*
1	10%
2	-5%
3	20%
4	15%

Your holding period return =

$$= (1 + r_1) \times (1 + r_2) \times (1 + r_3) \times (1 + r_4) - 1$$

$$= (1.10) \times (.95) \times (1.20) \times (1.15) - 1$$

$$= .4421 = 44.21\%$$

THAMMASAT
BUSINESS SCHOOL

Holding Period Return: Example

- An investor who held this investment would have actually realized an annual return of 9.58%:

Year	*Return*
1	10%
2	-5%
3	20%
4	15%

Geometric average return =

$$(1+r_g)^4 = (1+r_1)\times(1+r_2)\times(1+r_3)\times(1+r_4)$$

$$r_g = \sqrt[4]{(1.10)\times(.95)\times(1.20)\times(1.15)} - 1$$

$$= .095844 = 9.58\%$$

- So, our investor made 9.58% on his money for four years, realizing a holding period return of 44.21%

$$1.4421 = (1.095844)^4$$

Holding Period Return: Example

- Note that the geometric average is not the same thing as the arithmetic average:

Year	Return
1	10%
2	-5%
3	20%
4	15%

$$\text{Arithmetic average return} = \frac{r_1 + r_2 + r_3 + r_4}{4}$$

$$= \frac{10\% - 5\% + 20\% + 15\%}{4} = 10\%$$

THAMMASAT
BUSINESS SCHOOL

QUANTIFYING RISK AND UNCERTAINTY

THAMMASAT
BUSINESS SCHOOL

Risk and uncertainty

- Risk and uncertainty refers to randomness of events
- Risk is randomness in which events have measurable probabilities
- Uncertainty is randomness in which events have non-measurable probabilities because of imperfect information.
- “Risk, Uncertainty, and Profit” Frank Knight in 1921

THAMMASAT
BUSINESS SCHOOL

TYPES OF PROBABILITY

THAMMASAT
BUSINESS SCHOOL

Introduction

Fundamental Concepts

- A variable is **random** if its outcome is uncertain, where an **outcome** is an observable future value of the variable.
- An **event** is the specified set of possible outcomes of a random variable.
  - Events are **mutually exclusive** when the possible future outcomes can only occur one at a time and **exhaustive** when the set of outcomes includes every possible value the variable could take in the future.
  - Example: The future size of a dividend can be stated as a mutually exclusive and exhaustive event wherein dividends increase, decrease, or remain unchanged.
  - When the occurrence of one event does not affect the probability of the occurrence of another event, we say the events are **independent**.
  - Events that are not independent are **dependent**.

Probability

Probability is the fundamental building block of statistics.

- Probability is a number between 0 and 1 that describes the chance that a stated event from the set of possible outcomes will occur.
- A probability distribution is the set of probabilities and their associated outcomes that describes all possible outcomes and their associated probabilities.
- We typically use *P(E)* to denote the probability of event *E.*
  - Properties of probability

1. All probabilities must lie between 0 and 1: $$0 \leq P(E) \leq 1$$
2. For *n* mutually exclusive and exhaustive events, the sum of all probabilities must equal 1: $$\sum_{i=1}^{n} P(E_i) = 1$$

THAMMASAT
BUSINESS SCHOOL

Types of probability

Sources of probabilities

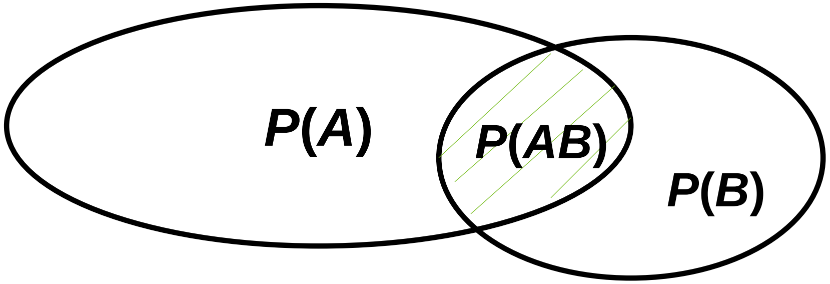
- In practice, we observe a number of different types of probability.
  - A **subjective probability** is a personal assessment of the likelihood of an event or set of events occurring in the future and is so named because it relies on the subjective judgment of the person making the assessment.
  - An **empirical probability** is one that is estimated from observed data, typically using the relative frequency at which an event or set of events has occurred in the past.
  - An **a priori probability** is one whose values are obtained from mathematical or logical analysis.

Conditional and unconditional probability

- When we observe the probability that an event (*A*) occurs without taking into account whether it is necessarily preceded by any other specific events, it is known as an **unconditional probability**.
  - Also known as a **marginal probability**.
  - Notation: $P(A)$
- When we observe the probability of a given event after taking into account that another event has already occurred (*A* occurs given that *B* has occurred), it is known as a **conditional probability**.
  - Notation: $P(A|B)$
  - Language: the probability of *A* given *B*

Joint Probability

- When two events occur, the combined probability of their occurrence is known as the joint probability.



  - Notation: $P(AB)$
  - Language: the probability of *A* and *B*
  - The calculation of a joint probability is governed by a set of probability rules, as are the calculations of the probability of other combinations of events.

Working with probabilities

- We use the **multiplication rule** to assess the joint probability of *A* and *B* occurring.

$$P(AB) = P(A|B)P(B)$$

  - Note that from this equation, we can calculate the conditional probability of *A* given *B,* as long as the probability of *B* isn't zero.

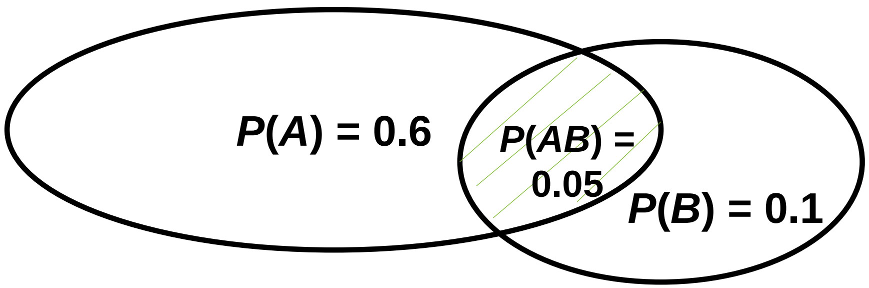
- We use the **addition rule** to assess the probability that *A* or *B* <u>or both</u> occur:

$$P(A \text{ or } B) = P(A) + P(B) - P(AB)$$

Addition rule

Focus on: Calculations

$$P(A\ or\ B) = P(A) + P(B) - P(AB)$$



- If the probability of relaxed import restrictions is 0.60 and the probability of a trade war is 0.10, then the probability of relaxed trade restrictions or a trade war is 0.65 when the joint probability of a trade war and relaxed trade restrictions is 0.05.

$$P(A\ or\ B) = 0.6 + 0.1 - 0.05 = 0.65$$

THAMMASAT BUSINESS SCHOOL

Multiplication rule

Focus on: Calculations

- Recall that our probability of relaxed trade restrictions has been estimated at 60%.
- If the probability of reduced sales, given that the trade restrictions are relaxed, is 80%, then the probability of relaxed trade restrictions and reduced sales is

$$P(AB) = P(A|B)P(B)$$

$$P(AB) = 0.8(0.6) = 0.48$$

THAMMASAT
BUSINESS SCHOOL

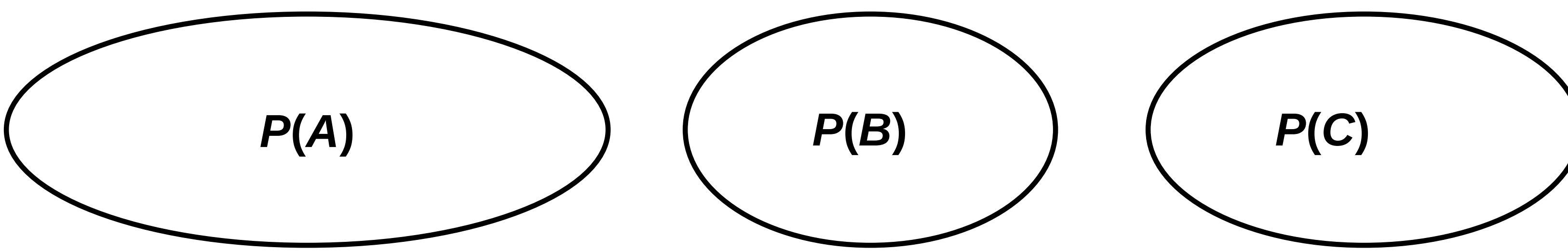
Dependent and Independent events

- When events are independent, the occurrence of one does not affect the probability of the other. In other words, *P*(*A*|*B*) = *P*(*A*) and *P*(*B*|*A*) = P(B).
  - To determine the joint probability, we use the multiplication rule:

$$P(AB) = P(A)P(B)$$

or, more generally,

$$P(ABC \ldots) = P(A)P(B)P(C) \ldots.$$



Joint probability of independent events

Focus on: Calculations

- Recall our analyst's estimated probability of relaxed import restrictions (60%).
- If *A* is the probability of a heavy winter snowfall (40%) and

  B is the independent probability of relaxed restrictions:
- Then the joint probability of relaxed import restrictions and a heavy winter snowfall is:

$$P(AB) = 0.4(0.6) = 0.24$$

THAMMASAT
BUSINESS SCHOOL

Self test

- Suppose in Japan, the probability that a person lives at least 80 years is 0.75 and probability that a person lives at least 90 years is 0.63. What is probability that a randomly selected 80 year old Japanese will survive to become 90?

Law of total probability

- Let B be an event with $P(B) > 0$ and $P(B') > 0$, then for any event A
- $P(A) = P(AB) + P(AB') = P(A|B)P(B) + P(A|B')P(B')$
- Ex. An insurance firm rents 35% of the cars for its customers from agency I and 65% from agency II. If 8% of the cars of agency I and 5% of the cars of agency II break down during rental periods, what is the probability that a car rented by this insurance firm breaks down.
- Ans: $P(A) = P(A|I)P(I) + P(A|II)P(II) = 0.08(0.35) + 0.05(0.65) = 0.0605$

Bayes' Rule

Bayes' Rule is often used to determine how a subjective belief should change given new evidence.

- First term: updated probability given new information, also known as posterior probability.
- Second term: probability of the new information given the event over the probability of the event.
- Third term: prior probability of the event.

$$P(\text{Event}|\text{Info}) = \left(\frac{P(\text{Info}|\text{Event})}{P(\text{Info})}\right) P(\text{Event})$$

THAMMASAT
BUSINESS SCHOOL

Define Event = Prob stock will go up in next 15 minutes of trade = P(U)

- You have been monitoring stock ABC for the past 20 days and taking information of order imbalance (OB) every 15 minutes.
- 20 days is equivalent to 360 observation intervals. Why?
- Of these intervals you count, $P(OB>0) = 0.6$. Denote $P(OB>0) = P(OB)$ for short.
- $P(OB>0|U) = 0.75$

- Today, the market has not opened. Your prior probability of $P(U) = 0.5$. Based on you historical (empirical) data collection, what is the probability that stock ABC will go up in first 15 minutes of market opening $P(U|OB)$?
- $P(U|OB) = [P(OB|U)*P(U)]/P(OB) = (0.75*0.5)/0.6 = 0.625$

THAMMASAT
BUSINESS SCHOOL

An experiment

- It is 9.50 am…
- What is the probability that CPN will close higher at end of mid-day trading session?
- Event = CPN closing higher than opening

What is your assessment of P(U|OIB) in next 15 minutes?

About SET | Rules / Regulations | Products / Services | Market Data | Companies / Securities Info | Related Entities | News

cpn | Search Symbol

Companies/Securities in Focus

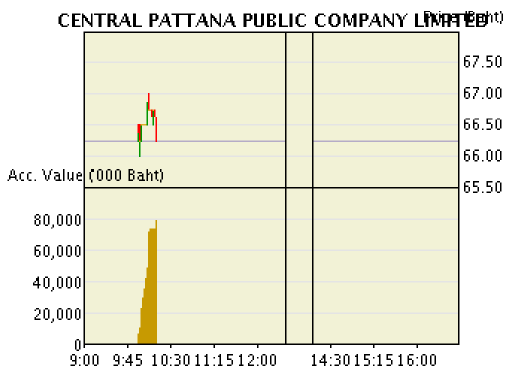
CPN : CENTRAL PATTANA PUBLIC COMPANY LIMITED

Choose Symbol: CPN

* Market data provided for educational purpose or personal use only, not intended for trading purpose.

- Profile
- F/S & Highlights
- Major Shareholder
- Rights & Benefits
- News
- Daily Quote
- Historical Trading

Acc. Value



Bid Price / Volume (Shares)	66.25 / 484,900
Offer Price / Volume (Shares)	66.50 / 417,400

ATO - At the open price
ATC - At the close price
Remark: Volume/value included volume/value from Auto Matching, Trade Report, and Odd Lot

* Last Update 23 Aug 2017 10:18:18
Market Status : Open(I)

Sign	
Last	**66.25**
Change	-
%Change	-
Prior	66.25
Open	66.25
High	67.00
Low	66.00
Volume (Shares)	1,370,311
Value ('000 Baht)	91,142.45
Average Price **	66.51
** Only Auto Matching	
Par (Baht)	0.50
Ceiling	86.00
Floor	46.50

Later…

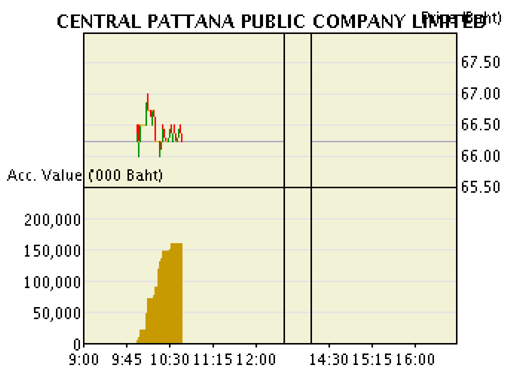
CPN : CENTRAL PATTANA PUBLIC COMPANY LIMITED

Choose Symbol CPN

* Market data provided for educational purpose or personal use only, not intended for trading purpose.

- Profile
- F/S & Highlights
- Major Shareholder
- Rights & Benefits
- News
- Daily Quote
- Historical Trading

Acc. Value



Bid Price / Volume (Shares)	66.25 / 339,100
Offer Price / Volume (Shares)	66.50 / 412,100

ATO - At the open price
ATC - At the close price
Remark: Volume/value included volume/value from Auto Matching, Trade Report, and Odd Lot

* Last Update 23 Aug 2017 10:45:09
Market Status : Open(I)

Sign	
Last	▲ 66.50
Change	+0.25
%Change	+0.38
Prior	66.25
Open	66.25
High	67.00
Low	66.00
Volume (Shares)	2,426,911
Value ('000 Baht)	161,172.27
Average Price **	66.41

** Only Auto Matching

Par (Baht)	0.50
Ceiling	86.00
Floor	46.50

Let's try another stock

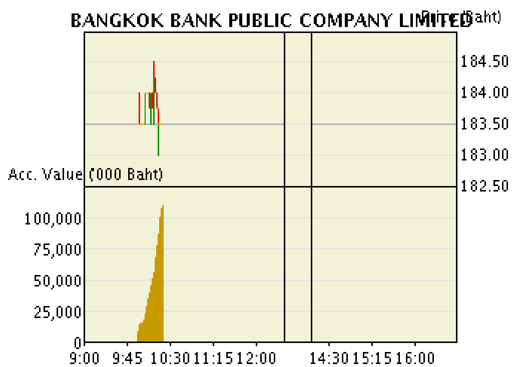
BBL : BANGKOK BANK PUBLIC COMPANY LIMITED

Choose Symbol BBL

* Market data provided for educational purpose or personal use only, not intended for trading purpose.

- Profile
- F/S & Highlights
- Major Shareholder
- Rights & Benefits
- News
- Daily Quote
- Historical Trading

Acc. Value



Bid Price / Volume (Shares)	183.50 / 21,800
Offer Price / Volume (Shares)	184.00 / 172,800

ATO - At the open price
ATC - At the close price
Remark: Volume/value included volume/value from Auto Matching, Trade Report, and Odd Lot

* Last Update 23 Aug 2017 10:27:04
Market Status : Open(I)

Sign	
Last	**183.50**
Change	-
%Change	-
Prior	183.50
Open	184.00
High	184.50
Low	183.00
Volume (Shares)	600,621
Value ('000 Baht)	110,406.41
Average Price **	183.82

** Only Auto Matching

Par (Baht)	10.00
Ceiling	238.00
Floor	128.50

Later…

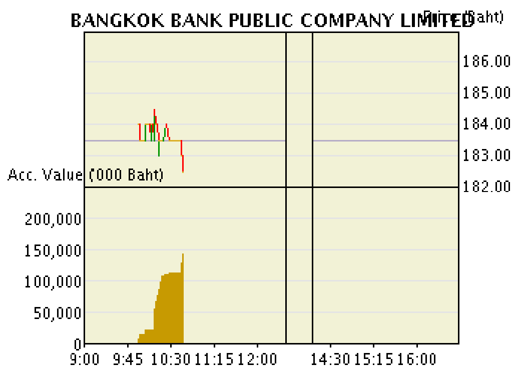
BBL : BANGKOK BANK PUBLIC COMPANY LIMITED

Choose Symbol: BBL

* Market data provided for educational purpose or personal use only, not intended for trading purpose.

- Profile
- F/S & Highlights
- Major Shareholder
- Rights & Benefits
- News
- Daily Quote
- Historical Trading

Acc. Value



Bid Price / Volume (Shares)	182.50 / 262,100
Offer Price / Volume (Shares)	183.00 / 139,400

ATO - At the open price
ATC - At the close price
Remark: Volume/value included volume/value from Auto Matching, Trade Report, and Odd Lot

* Last Update 23 Aug 2017 10:46:55
Market Status : Open(I)

Sign	
Last	▼ 182.50
Change	-1.00
%Change	-0.54
Prior	183.50
Open	184.00
High	184.50
Low	182.50
Volume (Shares)	787,931
Value ('000 Baht)	144,727.90
Average Price **	183.68
** Only Auto Matching	
Par (Baht)	10.00
Ceiling	238.00
Floor	128.50

Bayesian updating

Focus On: Calculations

- Suppose you have the following prior probabilities:
  - *P*(EPS exceeded consensus) = 0.45
  - *P*(EPS met consensus) = 0.30
  - *P*(EPS fell short of consensus) = 0.25
  - P(Expansion) = 0.41
- Given that DriveMed announces an expansion, what is the probability that prior quarter EPS (unreleased) exceeds consensus?

Bayesian updating

Focus On: Calculations

- You believe the conditional probabilities are
  - *P*(DriveMed expands | EPS exceeded consensus) = 0.75
  - *P*(DriveMed expands | EPS met consensus) = 0.20
  - *P*(DriveMed expands | EPS fell short of consensus) = 0.05

1. Use the total probability rule.
2. Apply Bayes' formula.

$$P(\text{Exceeds}|\text{Expands}) = \left(\frac{P(\text{Expands}|\text{Exceeds})}{P(\text{Expands})}\right) P(\text{Exceeds})$$

$$P(\text{Exceeds}|\text{Expands}) = \left(\frac{0.75}{0.41}\right) 0.45 = 0.823171$$

Bayesian updating

Focus On: Calculations

Prior Probabilities	Conditional Probabilities
P(EPS exceeded consensus) = 0.45	*P*(DriveMed expands | EPS exceeded consensus) = 0.75
P(EPS met consensus) = 0.30	*P*(DriveMed expands | EPS met consensus) = 0.20
P(EPS fell short of consensus) = 0.25	*P*(DriveMed expands | EPS fell short of consensus) = 0.05

Recall that you updated the probability that last quarter EPS exceeded the consensus from .45 to .823 after the expansion announcement.

$$P(\text{Exceeds}|\text{Expands}) = \left(\frac{0.75}{0.41}\right)0.45 = 0.823171$$

THAMMASAT
BUSINESS SCHOOL

Bayesian updating

Focus On: Calculations

- Update the prior probability that DriveMed's EPS met consensus.
  - Result is 0.146341, **down** from 0.3.
- Update the prior probability that DriveMed's EPS fell short of consensus.
  - Result is 0.030488, **down** from 0.25.
- Show that the three updated probabilities sum to 1.
  - Result: 0.030488 + 0.146341 + 0.823171 = 1

PROBABILITY DENSITY FUNCTIONS

Probability distribution

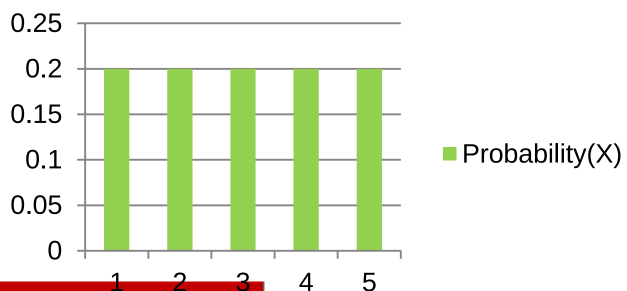
The set of probabilities for the possible outcomes of a random variable is called a "probability distribution."

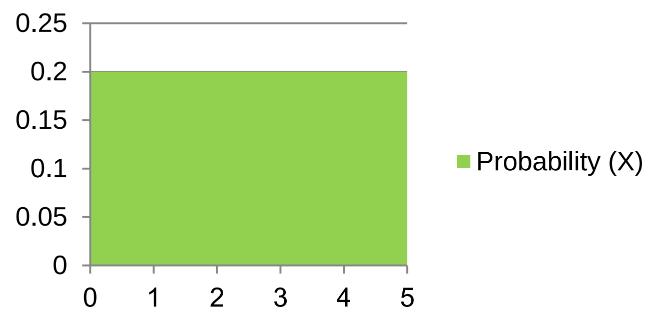
- The underlying foundation of most inferential statistical analysis is the concept of a probability distribution.
- The focus in the investments arena is on four probability distributions.
  1. Uniform
  2. Binomial
  3. Normal
  4. Lognormal (Coming in Financial Economics class)
- An understanding of probability distributions is critical to using such quantitative methods as hypothesis testing, regression, and time-series analysis.

Discrete and continuous random variables

A random variable is a variable whose future values are uncertain.

- **Discrete random variables** have a theoretically countable number of outcomes.
  - There may be an infinite number of them, but they are countable.
  - Price is a discrete random variable.
- **Continuous random** variables have a theoretically uncountable number of outcomes.
  - Rate of return is a continuous random variable.
  - Temperature is a continuous random variable.





Random Variables and outcomes

Focus on: Example of a Random Variable and Its Outcomes

- Consider a special dividend with five possible year-end values: $1, $5, $7, $10 and $11.
- Each of these values is known as an **outcome.**
- The random variable is described by its set of possible outcomes as
  - Div {$1, $5, $7, $10, $11}

Probability functions

The possible outcomes of a random variable and their associated probabilities are collectively known as a probability function.

- By convention, discrete random variable probability functions are denoted as *p(x)* and known as probability mass functions (pmf).
- Continuous random variable probability functions are denoted as *f(x)* and are known as probability density functions (pdf).
- Probability functions have two very important properties:

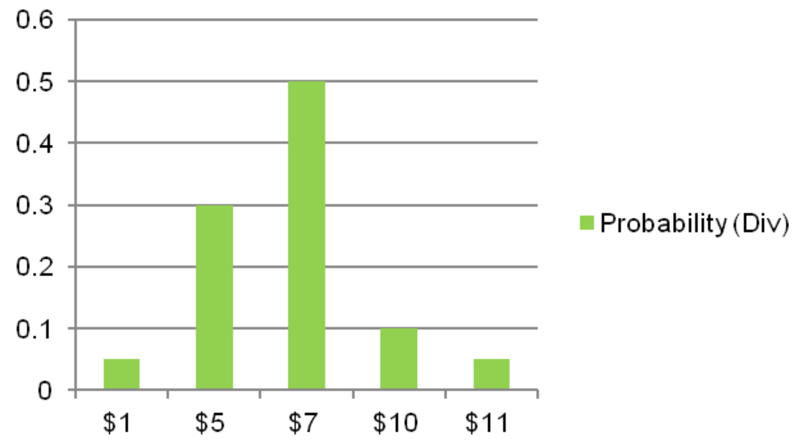
1. Any and all individual probabilities described by the probability function take on a value between 0 and 1 (including 0 and 1).
2. The sum of all probabilities described by the probability function is equal to 1.

$$0 \leq P(X_i) \leq 1 \qquad\qquad \sum_{i=1}^{n} P(X_i) = 1$$

Verifying a probability function

Focus on: Calculations

- Consider a special dividend with five possible year-end outcomes of \$1, \$5, \$7, \$10, and \$11.
- The probability of each outcome is 0.05, 0.3, 0.5, 0.1, and 0.05, respectively.
- Is this a valid probability function? It satisfies our two properties.



- This is a discrete random probability function.
  - The outcomes are countable, and it is a valid probability function.

THAMMASAT
BUSINESS SCHOOL

The probability density function (pdf)

The mathematical expression that describes the individual probabilities that a random variable will take on each of a set of specified values is known as its probability density function.

- For a discrete distribution, the pdf has discrete, countable, nonzero probabilities for every possible outcome.
- For a continuous distribution, the pdf has continuous, uncountable probabilities for each possible specified outcome in the set of infinite, uncountable outcomes. Hence, the probability of any specific outcome is zero.
  - For continuous distributions, this result means that the cumulative distribution function will be more useful and, to some extent, more meaningful.

The cumulative distribution function (cdf)

The mathematical expression that describes the probability that a random variable will be less than or equal to a specific value for all possible values of that variable is known as its cumulative distribution function.

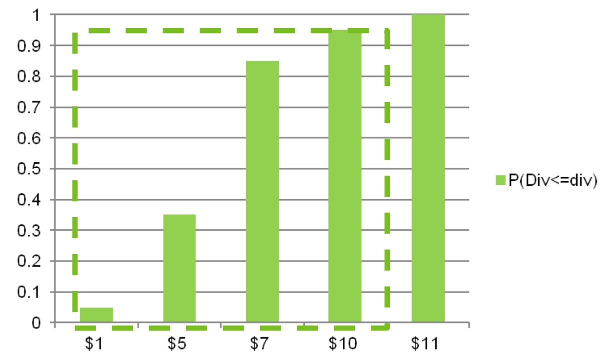
- The cumulative distribution function, denoted *F(x)*, is represented as the sum of the probabilities of the specified outcome and all prior outcomes in the distribution for each and every possible outcome.

- The cdf has the same properties as the pdf, in addition to
  1. All values of the cdf are between 0 and 1;
  2. As we increase the value of the specified outcome, the cdf must increase or remain constant.

The cdf in action

Focus on: Calculations

- Returning to the special dividend example, the cdf can be written and depicted as:



$$\text{cdf(Div)} = \begin{cases} P(\text{Div} \leq \$1) = 0.05 \\ P(\text{Div} \leq \$5) = 0.35 \\ P(\text{Div} \leq \$7) = 0.85 \\ P(\text{Div} \leq \$10) = 0.95 \\ P(\text{Div} \leq \$11) = 1.00 \end{cases}$$

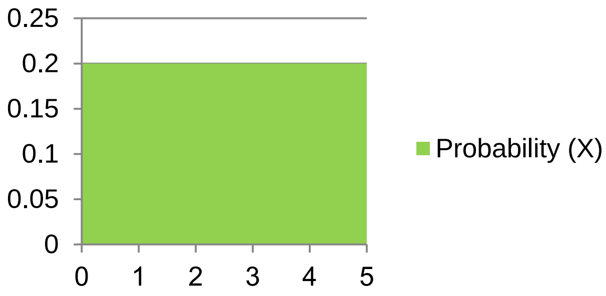
  - What is the probability of receiving at least a $10 dividend?
  - What is the probability of receiving more than a $7 dividend?

THAMMASAT
BUSINESS SCHOOL

The continuous uniform distribution

Recall that continuous variables are those whose possible outcomes cannot be counted.

- By analogy, the continuous uniform distribution is the continuous counterpart to the discrete uniform distribution.
  - This distribution is almost always the basis for generating random numbers in simulations; hence, it is a very important distribution.
- We generally use this distribution when we have no prior beliefs about the distribution of probabilities over outcomes (uncertainty in our beliefs) or when we believe probability is equally spread over the possible outcomes.



THAMMASAT
BUSINESS SCHOOL

The continuous uniform distribution

Focus On: Characterizing the Distribution

- The pdf and cdf for a continuous uniform distribution are written as

$$f(x) = \begin{cases} \dfrac{1}{b-a} \, when \; a < x < b \\ 0 \; otherwise \end{cases} \qquad F(x) = \begin{cases} 0 \; when \; x \leq a \\ \dfrac{x-a}{b-a} \; when \; a < x < b \\ 1 \; when \; x \geq b \end{cases}$$

- Probabilities are calculated from

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

- The mean and variance of a continuous uniformly distributed variable are

$$\mu = \frac{a+b}{2} \qquad \sigma^2 = \frac{(b-a)^2}{12}$$

THAMMASAT
BUSINESS SCHOOL

The continuous uniform distribution

Focus On: Calculations

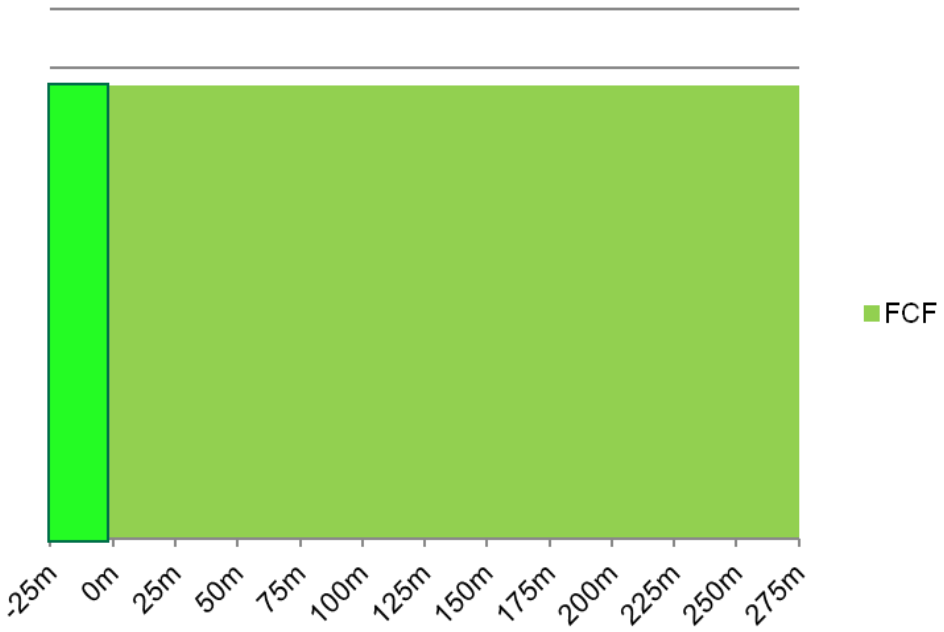
You are examining the forecasts of a buy-side analyst for the free cash flow (FCF) available to shareholders at a subject company. She estimates that the FCF will fall within –25m to 275m. You have decided to treat this as a continuous uniform variable.

What is the expected value of FCF?

- (–25 + 275)/2 = 125

What is the probability that FCF is negative?

- [0 – (–25)]/[275 – (–25)] = 0.083333



The normal distribution

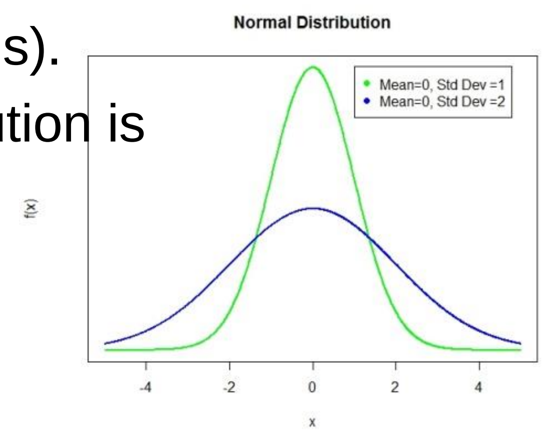
A continuous, symmetrical distribution that is compl described by its mean and variance.

- Arguably, it is the single most important distribution in statistics.
  - It plays a key role in modern portfolio theory and risk management.
  - The central limit theorem demonstrates that the sum and mean of a large number of independent random variables will be normally distributed even when the variables are not themselves normally distributed.
- The normal distribution ranges from infinitely negative to infinitely positive.
  - It is often used as a model for approximate returns.
- Linear combinations of normally distributed variables are also normally distributed.

The Normal distribution

A continuous, symmetrical distribution that is completely described by its mean and variance.

- Mean, median, and mode are equal.
- The normal distribution has skewness of zero.
  - Option returns are skewed; hence, they are not normally distributed.
- Kurtosis of 3 or excess kurtosis of 0 (3 – 3 = 0).
  - k > 3 → fat tails → underestimated probability of extreme values (the blue distribution has excess kurtosis).
  - This area is one for which the normal distribution is a poor approximation for stock returns, which have "fat tails."



The Normal distribution

Focus On: Characterizing the Normal Distribution

- The pdf for a normal distribution is written as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \text{ for } -\infty < x < \infty$$

- We indicate that a random variable is normally distributed as

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

- The mean and variance of a normally distributed variable are

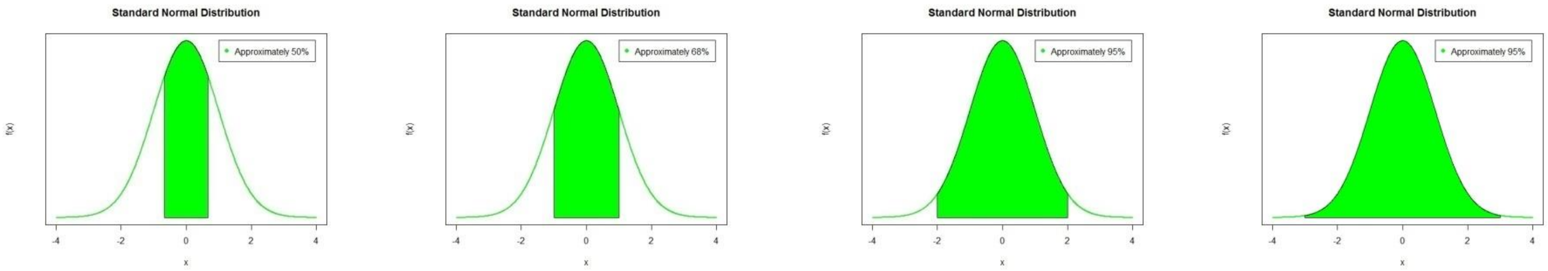
$$\mu \approx \bar{x} \qquad \sigma^2 = s^2$$

$$E(X) = \sum_{i=1}^{n} X_i P(X_i) \qquad \sigma^2(X) = \sum_{i=1}^{n} [X_i - E(X)]^2 P(X_i) = E\{[X - E(X)]^2\}$$

The normal distribution

Focus On: Confidence Intervals

- Approximately 50% of all observations fall in the interval μ ± (2/3)σ.
- Approximately 68% of all observations fall in the interval μ ± σ.
- Approximately 95% of all observations fall in the interval μ ± 2σ.
- Approximately 99% of all observations fall in the interval μ ± 3σ.



- We generally don't observe population mean and variance (μ and σ), but we can estimate them with sample mean and variance.
  - When we do, the same intervals apply, with the sample mean and variance used in place of their population analogs.

THAMMASAT
BUSINESS SCHOOL

The normal distribution

Focus On: Confidence Interval Calculations

- Your client's portfolio has a mean monthly return of 1.2% with a standard deviation of 3.7%. You assume for now that returns are normally distributed.
- Your client's return can be expected to fall in what range 50% of the time?
- 68% of the time?
- 95% of the time?

Expected value

Random Variable

- The expected value of a random variable is the probability-weighted average of the possible outcomes for that variable.

$$E(X) = \sum_{i=1}^{n} X_i P(X_i)$$

- We anticipate that there is a 15% chance that next year's return on holding Cleveland Corp will be 4%, a 60% chance it will be 6%, and a 25% chance it will be 8%. What is the expected return on Cleveland Corp stock?

$$E(X) = \sum_{i=1}^{n} 0.04(0.15) + 0.06(0.60) + 0.08(0.25)$$

$$E(X) = 0.062$$

THAMMASAT
BUSINESS SCHOOL

Variance

Random Variable

- The variance of a random value is the sum of the squared deviations from the expected value weighted by their associated probabilities.

$$\sigma^2(X) = \sum_{i=1}^{n} [X_i - E(X)]^2 P(X_i) = E\{[X - E(X)]^2\}$$

  - This value is a measure of the dispersion of possible values.
  - Because it has units that are squared, it is not easy to interpret. Accordingly, we use its positive square root, standard deviation, more often because it also measures dispersion but has the same units as expected value.
- The standard deviation of returns for Cleveland Corp. is then:

$$\sigma^2(X) = 0.15(0.04 - 0.062)^2 + 0.6(0.06 - 0.062)^2 + 0.25(0.08 - 0.062)^2$$
$$\sigma(R) = 0.01249$$

Covariance

Covariance and correlation are both measures of the extent to which two random variables move together.

- Covariance is the expected value of the product of each variable's deviation from its respective mean.

$$\sigma_{X,Y} = E\{[X - E(X)][Y - E(Y)]\}$$

$$\sigma_{R_i,R_j} = \sum_{i=1}^{n} P(R_i)[R_i - E(R_i)][R_j - E(R_j)]$$

THAMMASAT
BUSINESS SCHOOL

Correlation

Covariance and correlation are both measures of the extent to which two random variables move together.

- Correlation is a scaled transformation of covariance wherein the extent of comovement is measured along a scale from exactly the same movement in the same direction to exactly the same movement in opposite directions.
  - When two variables move the same degree in opposing directions, they are said to be perfectly negatively correlated.
  - When two variables move the same degree in the same direction, they are said to be perfectly positively correlated.
  - When there is absolutely no commonality of movement, the variables are said to be uncorrelated.

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

$$-1 \leq \rho_{X,Y} \leq 1$$

Covariance and Correlation

Focus On: Calculations

- We anticipate that there is a 15% chance that next year's stock returns for Cleveland Corp. will be 4%, a 60% chance they will be 6%, and a 25% chance they will be 8%. The standard deviation of returns is 1.249%, and the expected value is 6.2%.
- We anticipate that the same probabilities and states are associated with a 2% return for High Noon Inc., a 3% return, and a 3.5% return. The standard deviation of High Noon Inc. returns is then 0.46%, and its expected value is 2.975%.
  - What is the covariance between Cleveland and High Noon returns?
  - What is the correlation between Cleveland and High Noon returns?

$$\sigma_{R_i,R_j}$$
$$= 0.15(0.04 - 0.062)(0.02 - 0.02975)$$
$$+0.6(0.06 - 0.062)(0.03 - 0.02975)$$
$$+0.25(0.08 - 0.062)(0.035 - 0.02975)$$
$$= 0.0000555$$

$$\rho_{R_i,R_j} = \frac{0.0000555}{0.0046(0.01249)} = 0.966$$

THAMMASAT
BUSINESS SCHOOL

Standard Normal

A normal distribution with a mean of 0 and standard deviation of 1 is called "standard normal."

- The prevalence of the normal distribution has led to a process whereby probability tables that have been calculated for a standard normal distribution can be used to make probability statements for any normally distributed variable.
- This process is known as "standardizing" and is accomplished by:

1. Taking the observation(s) of interest and subtracting the mean of that observation's observed distribution;
2. Dividing the result by the observed distribution's standard deviation.

$$Z = \frac{X - \mu}{\sigma} \qquad z = \frac{x - \bar{x}}{s}$$

THAMMASAT BUSINESS SCHOOL

Standard normal

- The pdf for standard normal is

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

- If X ~ N(μ, σ^2) The mean and variance of the standardized X is N(0,1)

Standard normal

- The scores on an achievement test given to 100,000 students are normally distributed with mean 500 and SD 100. What should the score of a student be to place him among the top 10% of all students?

- Need to find P(X>=x) = 0.10

- This is the same as P(X < x) = 0.90

$$P\left(\frac{X-500}{100} < \frac{x-500}{100}\right) = 0.90$$

$$\Phi\left(\frac{x-500}{100}\right) = 0.90$$

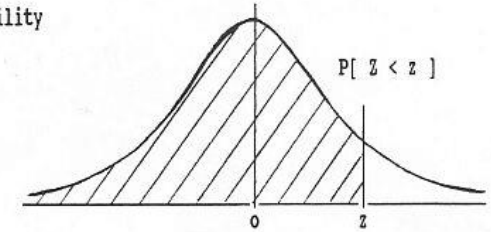
- From next slide, we have $\Phi(1.28) = 0.8997$, implying (x-500)/100 = 1.28; therefore x approx. 628

STANDARD STATISTICAL TABLES

1. Areas under the Normal Distribution

The table gives the cumulative probability up to the standardised normal value z
i.e.

$$P[\ Z < z\ ] = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}Z^2)\ dZ$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5159	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7854
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	[illegible]	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9865	0.9868	0.9871	0.9874	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9924	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9980	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

z	3.00	3.10	3.20	3.30	3.40	3.50	3.60	3.70	3.80	3.90
P	0.9986	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

The Standard normal distribution

Focus On: Calculations

- Your client’s portfolio has a mean monthly return of 1.2% with a standard deviation of 3.7%. You assume for now that returns are normally distributed.
  - What is the chance that returns will be between –2.5% and 4.9%?
  - What is the chance that returns will be negative?
- A stop-loss order automatically sells the stock if the price is below a set amount. You can set a stop-loss so that the portfolio is liquidated when it is triggered. How often will such a stop-loss be triggered if you set it so that it triggers when losses are below 1%?