Study Session # 2, Reading # 5

"The Time Value of Money"

Revision-1, released on November 23, 2010.

Compound Interest or Interest on Interest

Growth in the value of investment includes, interest earned on:

- Original principal.
- Previous period's interest earnings.

Time Line

Diagram of the cash flows associated with a TVM problem.

Discounting

Moving CF to the beginning of an investment period to calculate PV.

$$PV = \frac{FV}{(1+i)^N}$$

$\frac{1}{(1+i)^N}$ is PV factor

Compounding

Moving cash flow to the end of the investment period to calculate FV.

$FV = PV\ (1+i)^N$

$(1+i)^N$ is FV factor

Required interest rate on security. = Nominal RFR. + Default risk premium. + Liquidity premium. + Maturity risk premium.

- Nominal RFR. → Real RFR + Expected inflation rate.
  - Real RFR → It shows an increase in purchasing power.
- Default risk premium. ⇓ Premium for the risk that borrower will not make the promised payments in a timely manner.
- Liquidity premium. ⇓ Premium for receiving less than fair value for an investment if it must be sold quickly.
- Maturity risk premium. ⇓ Longer-term bonds have more maturity risk, because their prices are more volatile.

Loan Amortization

Process of paying off a loan with a series of periodic loan payments, whereby a portion of the outstanding loan amount is paid off, or amortized, with each payment.

Perpetuity

- Perpetual annuity.
- Fixed payment at set intervals over an infinite time period.
- $\frac{1}{r}$ is the discounting factor for perpetuity.

Annuity

Stream of equal cash flows accruing at equal intervals.

⇓

Two types

⇐ **Annuity Due**: Cash flows occur at the beginning of each period.

⇒ **Ordinary Annuity**: Cash flows occur at the end of each period.

Cash flow Additivity Principle

PV of any stream of cash flows equals the sum of PV of each cash flow.

PV of annuity due. > PV of ordinary annuity.

Interpretations of Interest Rate

- Required rate of return.
- Discount rate.
- Opportunity cost.

Effective Annual Rate (EAR)

- Rate of return actually being earned after adjustments have been made for different compounding periods.
- $EAR = (1+ \text{periodic rate})^m - 1$
- Stated rate will be equal to the actual (effective) rate only when it is compounded annually.

Study Session # 2, Reading # 6

"Discounted Cash Flow Applications"

D.R. = Discount Rate.

NPV:

- PV of expected cash inflows. – PV of expected cash outflows.
- NPV = $\sum_{t=0}^{N} \frac{CF_t}{(1+r)^t}$
- D.R used is the market based opportunity cost of capital.
- NPV assumes reinvestment at D.R.

Decision rule:

NPV	Decision	Impact
+ve (IRR> r)	Accept	Increase shareholder's wealth.
0	-	Size of the company rises but shareholder's wealth is not affected.
-ve (IRR< r)	Reject	Decreases shareholder's wealth.

IRR

- The D.R. at which NPV = 0.
- The rate of return at which; PV inflows = PV outflows.
- It assumes reinvestment at IRR.

Decision rule

- IRR > r ⇒ Accept
- IRR < r ⇒ Reject

For Single Project: IRR / NPV rules lead to exactly the same accept /reject decision

For Mutually Exclusive Project: Select the project with the greatest NPV.

Problems in IRR

For mutually exclusive projects, NPV & IRR may give conflicting project rankings due to:

- Different sizes of project's initial cost
- Differences in timings of cash flows.

Money-Weighted Return (MWR)

- IRR of an investment.
- It is highly sensitive to the timing & amount of withdrawals from & additions to a portfolio.
- If one has complete control over the timings of cash flows then it is more appropriate.

Time-Weighted Return (TWR)

- It measures compound growth (Geometric return).
- It is not affected by the timings of cash flows.
- It is preferred over MWR.
- TWR can't be calculated if we don't know the period end values of investment.

Funds contributed prior to a period of relatively..

- Poor returns → MWR< TWR
- High returns → MWR> TWR

Money Market Yield (CD equivalent yield) r_{MM}

- Annualized HPY.
- Assumes 360-day year.
- Does not incorporate effects of compounding.
- $r_{MM} = HPY \times \frac{360}{t}$

Effective Annual Yield (EAY)

- Annualized HPY.
- Assumes 365-day year.
- Incorporates effects of compounding.
- $EAY = (1+HPY)^{365/t} - 1$

Bank Discount Yield (r_{BD})

- Dollar discount from the face (par) value as a fraction of the face value.
- Based on face value.
- Not based on market or purchase price.
- $r_{BD} = \frac{D}{F} \times \frac{360}{t}$

Converting r_{BD} into r_{MM}

$$r_{MM} = \frac{360 \times r_{BD}}{360 - (t \times r_{BD})}$$

Bond –equivalent yield (BEY)

BEY = 2 × (semi annual effective yield.)

Holding Period Yield or Holding Period Return (HPR)

- Total return an investor earns between the purchase date & the sale or maturity date.

$$HPY \text{ or } HPR = \frac{P_1 - P_0 + D_1}{P_0} = \frac{P_1 + D_1}{P_0} - 1$$

- It is the actual return an investor receives.

"Statistical Concepts & Market Returns"

Population
Set of all possible members.

Sample
Subset of population.

Statistics
Refers to data & methods used to analyze data.

Parameter
Measures characteristics of population.

Sample Statistic
Measures characteristics of a sample.

Two Categories

Descriptive Statistics
Used to summarize & consolidate large data sets into useful information.

Inferential Statistics
Forecasting, estimating or making judgment about a large set based on a smaller set.

Types of Measurement Scales

Nominal Scale
- Least accurate.
- No particular order or rank.
- Provides least info.
- Least refined.

Ordinal Scale
- Provides ranks/orders.
- No equal difference b/w scale values.

Interval Scale
- Provides ranks/orders.
- Difference b/w the scales are equal.
- Zero does not mean total absence.

Ratio Scale
- Provides ranks/orders
- Equal differences b/w scale.
- A true zero point exists as the origin.
- Most refined.

Frequency Distribution
Tabular (summarized) presentation of statistical data.

Constructing a Frequency Distribution

1. Define Intervals / Classes
- Interval is set of values that an observation may take on.
- Intervals must be,
  - All-inclusive.
  - Non-overlapping.
  - Mutually Exclusive.

Importance of Number of Intervals

Too few intervals. → Important characteristics may be lost.

Too many intervals. → Data may not be summarized well enough.

2. Tally the observations
Assigning observations to their appropriate intervals.

3. Count the observations
Count actual number of observations in each interval i.e., absolute frequency of interval.

Modal Interval
Interval with highest frequency.

Cumulative Absolute Frequency
Sum of absolute frequencies starting with lowest interval & progressing through the highest.

Relative Frequency
%age of total observations falling in each interval.

Cumulative Relative Frequency
Sum of relative frequencies starting with the lowest interval & progressing through highest.

Histogram
- Bar chart of continuous data frequency distribution.
- Helps in quickly identifying the modal interval.
- X-axis: Class intervals.
- Y-axis: Absolute frequencies.

Frequency Polygon
- X-axis: Mid points of each interval.
- Y-axis: Absolute frequencies.

Measures of Central Tendency

- Identify centre of data set.
- Measure of reward.
- Used to represent typical or expected values in data set.

Mean

- Sum of all values divided by total number of values.
- Population = $\frac{\Sigma X}{N} = \mu$
- Sample = $\frac{\Sigma x}{n} = \bar{x}$

Properties:

- Mean includes all values of data set.
- Mean is unique for each data.
- All intervals & ratio data sets have a Mean.
- Sum of deviations from Mean is always zero i.e., $\Sigma(x - \bar{x}) = 0$
- Mean is the best estimate of true mean & the value of next observation.

Shortcoming:

- Mean is affected by extremely large & small values.

Median

- Midpoint of an arranged data set.
- Divides data into two equal halves.
- It is not affected by extreme values; hence it is a better measure of central tendency in the presence of extremely large or small values.

Weighted Mean

It recognizes the disproportionate influence of different observations on mean.

$$\bar{X}_w = \sum_{i=1}^{n} wi\ Xi; \quad \sum_{1=i}^{n} wi = 1$$

Mode

- Most frequent value in the data set.

No. of Modes	Names of Distributions
One →	Unimodal
Two →	Bimodal
Three →	Trimodal

Geometric Mean (GM)

- Calculating multi periods return.
- Measuring compound growth rates.

$$G = \sqrt[n]{X1 \times X2\ \ldots \times Xn}$$

(applicable only to non-negative values)

$$1 + R_G = \sqrt[n]{(1+R_1) \times (1+R_2) \times \ldots \times (1+R_n)}$$

Harmonic Mean (H.M)

$$H.M = \frac{N}{\Sigma \left[\frac{1}{X}\right]}$$

H.M is used:

- When time is involved.
- Equal $ investment at different times.

For values that are not all equal

H.M < GM < AM

Quantiles

Quantiles: $L_y = [n-1]\frac{y}{100}$

Quartiles: Distribution divided into 4 parts (quarters).
Quintiles: Distribution divided into 5 parts.
Deciles: Distribution divided into 10 parts.
Percentiles: Distribution divided into 100 parts (percents).

Measures of Location ⇒ Measures of Central Tendency + Quantiles

Dispersion

- Variability around the central tendency.
- Measure of risk.

Range

Max Value — Min Value

Population Variance 'σ^2'

Averaged squared deviations from mean.

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{N}$$

Population Standard Deviation (S.D) 'σ'.

Square root of population variance.

$$\sigma = \sqrt{\frac{\Sigma(x-\mu)^2}{N}}$$

Mean Absolute Deviation (MAD)

Average of absolute deviations from mean:

$$MAD = \frac{\Sigma|X - \bar{X}|}{n}$$

Sample Variance

$$s^2 = \frac{\Sigma(X - \bar{X})^2}{n-1}$$

Sample Standard Deviation

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n-1}}$$

⇒

Using 'n-1' observations

Using entire number of observations 'n' will systemically underestimate the population parameter & cause the sample variance & S.D to be referred to as biased estimator.

In general S.D > MAD

Relative Dispersion

Amount of variability relative to a reference point.

Coefficient of Variation

- CV= $\frac{S_x}{\bar{X}}$ i.e., risk per unit of expected return.
- Helps make direct comparisons of dispersion across different data sets.

Sharpe Ratio

- Measures excess return per unit of risk.
- Sharpe ratio = $\frac{\bar{r}_p - r_f}{\sigma_p}$
- Higher sharpe ratios are preferred.

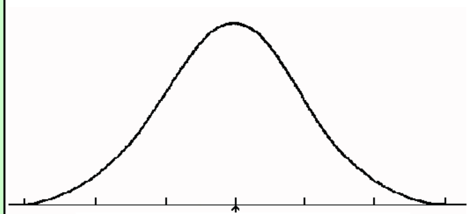
Chebyshev's Inequality

Gives the %age of observations that lie within 'k' standard deviations of the mean is at least $1-\frac{1}{k^2}$ for all k>1, regardless of the shape of the distribution.

$\pm$ 1.25 s.d	$\Rightarrow$	36% obs.
$\pm$ 1.5 s.d	$\Rightarrow$	56% obs
$\pm$ 2 s.d	$\Rightarrow$	75% obs.
$\pm$ 3 s.d	$\Rightarrow$	89% obs.
$\pm$ 4 s.d	$\Rightarrow$	94% obs.

Symmetrical Distribution

- Identical on both sides of the mean.
- Intervals of losses & gains exhibit the same frequency.
- Mean = Median = Mode.



Mean = Median = Mode.

Skewness

Non Symmetrical.

Sample Skewness

- Sum of cubed deviations from mean divided by number of observations & cubed standard deviation.

$$s_k = \frac{1}{n}\frac{\sum(x-\bar{x})^3}{s^3}$$

- $|S_k|$ >0.5 is considered significant level of skewnes.

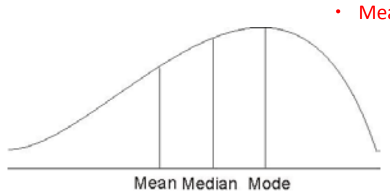
Negatively Skewed

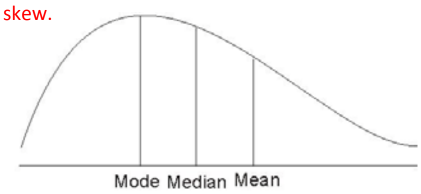
- Longer tail towards left.
- More outliers in the lower region.
- More – ve deviations.
- Mean < Median < Mode

Positively Skewed

- Longer tail towards right.
- More outliers in the upper region.
- More + ve deviations.
- Mean > Median > Mode.

Hint
- Median is always in the center.
- Mean is in the direction of skew.





Kurtosis

- Measure of degree of more or less peaked than a normal distribution.
- Kurtosis of normal distribution is 3.
- Excess kurtosis = sample kurtosis-3

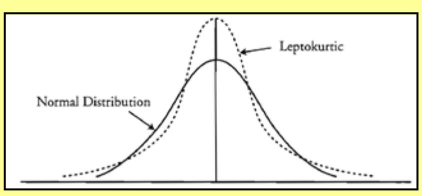
$$where, sample\ kurtosis = \frac{1}{n}\frac{\Sigma(X-\bar{X})^4}{s^4}$$

- Excess kurtosis value exceeding absolute 1 is considered large.

Distribution		Excess Kurtosis
Leptokurtic	$\Rightarrow$	>0
Mesokurtic (Normal)	$\Rightarrow$	=0
Platykurtic	$\Rightarrow$	<0

In Leptokurtic distribution, there is a greater probability of both, very small & very large deviation sfrom mean.

Greater +ve kurtosis & more – ve skew. $\Rightarrow$ Increased Risk.



"Probability Concepts"

Random Variable
Uncertain & unforcastable number.

Outcome
An observed value of a random variable.

Event
A single outcome or a set of outcomes.

Mutually Exclusive Events
Both can't happen at the same time.
P(A/B) = 0 &
P(AB) = P(A/B). P(B) =0

Exhaustive Events
Include all possible outcomes.

Two Defining Properties of Probability

- $0 \leq P(E) \leq 1$
  i.e., Probability of an event lies b/w 0 & 1.
- $\Sigma P(E_i) = 1$
  i.e., Total probability is equal to 1.

Probability in terms of

- **Odds for the event**: Probability of occurrence divided by probability of non-occurrence.
- **Odds against the event**: Probability of non-occurrence divided by probability of occurrence.

Probability

Empirical Probability
- Based on historical facts or data.
- No judgments involved.
- Historical + non random.

A Priori Probability
- Based on sure logic or formula.
- Random + historical.

Subjective Probability
- An informal guess.
- Involves personal judgment.

Objective Probability (Empirical Probability and A Priori Probability)

Unconditional Probability
- Marginal probability.
- Probability of occurrence of an event- regardless of the past or future occurrence.

Conditional Probability; P(A/B)
- Probability of the occurrence of an event is affectedly the occurrence of another event.
- It is also known as likelihood of an occurrence.
- '/' denotes 'given' or 'conditional' upon.
- $P(A/B) = \frac{P(AB)}{P(B)}$
- Mutually exclusive events P(A/B) = 0.
- For independent events, P(A/B) = P(A)

Independent Events
- Events for which occurrence of one has no effect on occurrence of the other.
- P(A/B) = P(A)
- P(B/A) = P(B)

Multiplication Rule (Joint Probability)
- Probability that both events will occur.
- P(AB) = P(A/B). P(B)
  ⇒ For mutually exclusive events;.
  P(A/B) = 0, hence,
  P(AB) = 0.

Addition Rule
- Probability that at least one event will occur.
- P(A or B) = P(A) + P(B) - P(AB)
  ⇒ For mutually exclusive events.
  P(A or B) = P(A) + P(B).

Total Probability Rule
It highlights the relationship b/w unconditional & conditional probabilities of mutually exclusive & exhaustive events.
$P(R) = P(RI) + P(RI^c)$
$= P(R/I). P(I) + P(R/I^c). P(I^c)$

Expected Value
- Weighted avg. value of a random variable that results from multiple experiments.
- It is the best guess of the outcome of a random variable.

Conditional Expected Value
- Calculated using conditional probabilities.
- Are contingent upon the occurrence of some other event.

Study Session # 2, Reading # 8

Covariance

- Measure of how two assets move together.
- It measures only direction.
- $-\infty \le Cov(x, y) \le +\infty$ (property).
- It is measured in squared units.
- $Cov(R_i,R_j) = E\{[R_i - E(R_i)]\,[R_j - E(R_j)]\}$
  $= \sum P(S)\,[R_i - E(R_i)]\,[R_j - E(R_j)]$.
- $Cov(R_A,R_A) = variance(R_A)$ (property).

Covariance	Variables tend to
+ ve ⇒	Move in same direction.
- ve ⇒	Move in opposite direction.
'0' ⇒	No linear relationship.

Correlation

- Measures the direction as well as the magnitude.
- It is a standardized measure of co-movement.
- It has no units.
- $-1 \le corr(Ri,Rj) \le +1$.

Value	Correlation	Variables tend to
+1 ⇒	Perfectly positive ⇒	Move proportionally in the same direction.
-1 ⇒	Perfectly negative ⇒	Move proportionally in the opposite direction.
0 ⇒	Uncorrelated ⇒	No linear relationship.

$$Corr(R_i,R_j) = \frac{Cov(R_i,R_j)}{\sigma(R_i)\,\sigma(R_j)}$$

Portfolio

Expected Value

$$E(R_p) = \sum_{i=1}^{N} w_i E(R_i)$$

Variance

$$Var(R_p) = \sum_{i=1}^{N} w_i^2 \sigma_i^2 + \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j Cov_{ij}$$

⇒ Where $w_i = \dfrac{\text{market value of investment in asset 'i'}}{\text{market value of the portfolio}}$

Baye's Formula

⇒ Used to update a given set of prior probabilities in response to the arrival of new information.

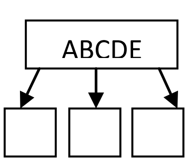
$$\text{Updated Probability} = \frac{\text{probability of new info.}}{\text{unconditional probability of new info.}} \times \text{prior probability of event.}$$

Counting Methods

Labeling Formula

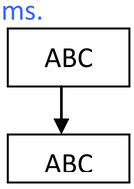
$$\frac{n!}{n_1! \ldots n_k!}$$

Assigning each element of the entire group in one of the three or more subgroups.



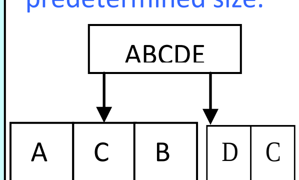
Factorial [!]

- Arranging a given set of 'n' items.
- No subgroups.
- There are n! ways of arranging 'n' items.



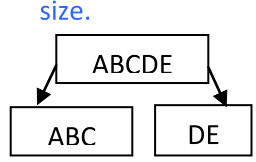
Permutation [^nP_r]

Specific ordering of a group of objects into only two groups of predetermined size.



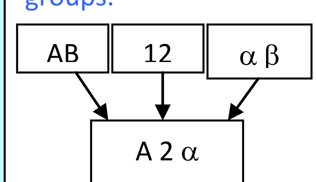
Combination [^nC_r]

- Choosing 'r' items from a set of 'n' items when order does not matter.
- It applies to only two groups of predetermined size.



Multiplication Rule

- Selecting only one item from each of the two or more groups.



"Common Probability Distributions"

Probability Distribution

- Describes the probabilities of all possible outcomes for a random variable.
- Sum of probabilities of all possible outcomes is 1.

Probability Function

Probability of a random variable being equal to a specific value.

Properties:

- $0 \le p(x) \le 1$
- $\Sigma\, p(x) = 1$

	Discrete	Continuous
Random Variable	Finite (measurable) # of possible outcomes.	Infinite (immeasurable) # of possible outcomes.
Distribution	• P(x) can't be 0 if 'x' can occur. • We can find the probability of a specific point in time.	• P(x) can be zero even if 'x' can occur. • We can't find the probability of a specific point in time.

Probability Density Function (PDF)

- It is used for continuous distribution.
- Denoted by f(x).
- Finds the probability of an outcome within a particular range (b/w two values).
- Probability of any one particular outcome is zero.

Cumulative Distribution Function (CDF)

- Calculates the probability of a random variable 'x' taking on the value less than or equal to a specific value 'x'.
- $F(x) = P(X \le x)$

Binomial Distribution

Properties:

- Two outcomes (success & failure).
- 'n' number of independent trials.
- With replacement.
- Probability of success remains constant.
- $E(x) = np$.
- $p(x) =$

$$\frac{n!}{(n-x)!.x!}\, p^x (1-p)^{n-x}$$

$$= {}^{n}C_x . P^x\, q^{n-x}$$

Binominal Tree

- Shows all possible combinations of up & down moves over a number of successive periods.
- Node: Each of the possible values along the tree.
- U is up-move factor.
- D is down-move factor (1/U).
- p is probability of up move.
- (1-p) is probability of down move.

Discrete uniform random variable ⇒ All outcomes have the same probability.

Uniform Probability Distribution

Discrete

- Has a finite number of specified outcomes.
- $P(x) \times k$. K is the probability for 'k' number of possible outcomes in a range.
- cdf: $F(x_n) = n.p(x)$.

Continuous

- Defined over a range with parameters 'b' (upper limit) & 'a' (lower limit).
- cdf: It is linear over the variable's range.
- Properties:
- $P(x < a \text{ or } x > b) = 0$
- $P(x_1 \le X \le x_2) = \frac{x_2 - x_1}{b - a}$
- For all $a \le x_1 < x_2 \le b$.

Univariate Distribution

Probability distribution of a single random variable.

Multivariate Distribution

Specifies the probabilities associated with a group of random variables.

Normal Distribution

- Completely described by μ & σ^2.
- Stated as $X \sim N(\mu, \sigma^2)$.
- Symmetric about its mean, Skewness =0. $P(X \le \mu) = 0.5 = P(X \ge \mu)$.
- Kurtosis = 3.
- Linear combination of normally distributed random variables is also normally distributed.
- Large deviations from mean are less likely than small deviations.
- Probability becomes smaller & smaller as we move away from mean, but never becomes '0'.

Multivariate Normal Distribution

It can be completely described by three parameters

- n means.
- n variances.
- $\frac{n(n-1)}{2}$ pair wise correlations.

Log-Normal Distribution

- Generated by function e^x where 'x' is normally distributed.
- Skewed to the right.
- Bounded from below by 0.

Standard Normal Distribution

- Has $\mu = 0$ & $\sigma = 1$ i.e., $N \sim (0,1)$.
- Standardization: Process of converting an observed value of a random variable to its z-value.
- $z = \frac{x-\mu}{\sigma}$

Confidence Interval

Range of values around the expected value within which actual outcome is expected to be some specified percentage of time.

Confidence Interval	%age
x ± 1s	68%
x ± 1.65s	90%
x ± 1.96s	95%
x ± 2s	95.45%
x ± 2.58s	99%
x ± 3s	99.73%

Compounded Rates of Returns

Discrete

Daily, annually, weekly, monthly compounding

Continuous

- $\ln(S_1/S_0) = \ln(1+HPR)$
- These are additive for multiple periods.
- Effective annual rate based on continuous compounding is given as: $EAR = e^{Rcc}-1$

Historical Simulation

- Based on actual values & actual distributions of the factors i.e., based on historical data.

Limitation:

- History does not repeat itself.
- Historical data does not provide flexibility.

Roy's Safety First Criterion

- Optimal portfolio minimizes the probability that the return of the portfolio falls below some minimum acceptable level.
- Minimize $P(R_P < R_L)$.
- SFR is the number of standard deviations below the mean.
- SFRatio = $\frac{[E(R_P) - R_L]}{\sigma_P}$
- Choose the portfolio with greatest SFRatio.

Shortfall Risk

Probability that portfolio value will fall below some minimum level at a future date.

Monte-Carlo Simulation

- Repeated generation of one or more factors (e.g. risk) that affect required value (e.g., stock price) in order to generate a distribution of the values (stock price).
- We have the flexibility of providing the data.

Simulation Procedure for Stock Option Valuation

Specify prob. dist. of stock prices & relevant interest rate as well as their parameters. → Randomly generate values of stock prices & interest rates. → Value the options for each pair of risk factors. → Calculate mean option value performing many iterations & use it as estimated option value.

Uses

- Valuing complex securities.
- Simulating gains / losses from trading strategy.
- Estimating value at risk (VAR).
- Examining variability of the difference b/w assets & liabilities of pension funds.
- Valuing portfolio with non - normal return distribution.

Limitations

- Complex procedure.
- Highly dependent on assumed distributions.
- Based on a statistical rather than an analytical method.

Study Session # 3, Reading # 10

"Sampling & Estimation"

s.e = standard error
↑ = rises

→ = approaches to
d.f = degrees of freedom
n = sample size

Sample
A subgroup of population.

Sample Statistic
- It describes the characteristic of a sample.
- Sample statistic itself is a random variable.

Methods of Sampling

Simple Random Sampling
- Each item of the population under study has equal probability of being selected.
- There is no guarantee of selection of items of a particular category.

Stratified Random Sampling
- Uses a classification system.
- Separates the population into strata (small groups) based on one or more distinguishing characteristics.
- Take random sample for each stratum.
- It guarantees the selection of items from a particular category.

Systematic Sampling
- Select every n^{th} number.
- It gives approximately random sampling.

Sampling error
Sample – Corresponding Statistic Population Parameter.

Sampling Distribution
Probability distribution of all possible sample statistics computed from a set of equal size samples randomly drawn.

Standard Error (s.e) of Sample Mean
- Standard deviation of the distribution of sample means.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

- If σ is not known then; $s_{\bar{x}} = \frac{s}{\sqrt{n}}$
- As n ↑; $\bar{x}$ approaches μ and s.e ↓.

Data	Time	Variables
Time series	Observations taken over equally spaced time interval	Single
Cross- sectional	Single point estimate	Different variables

Data	Entity	Characteristics
Longitudinal	Same	Multiple
Panel	Multiple	Same

Central Limit Theorem (CLT)
For a random sample of size 'n' with;
- population mean μ,
- finite variance σ^2, the sampling distribution of sample mean $\bar{x}$ approaches a normal probability distribution with mean 'μ'& variance as 'n' becomes large. i.e, As n ↑; $\bar{x}$ ~

Properties of CLT
- For n ≥ 30 ⇒ sampling distribution of mean is approx. normal.
- For n ≤ 30 ⇒ Not normal.
- Mean of distribution of all possible samples = population mean 'μ'.
  - Variance o distribution = σ^2

CLT applies only when samples is random.

Point Estimate (P.E)
- Single (sample) value used to estimate population parameter.

$$\bar{X} = \frac{\Sigma X}{n}$$

Estimator: Formula used to compute P.E.

Confidence Interval (CI) Estimates
- Results in a range of values within which actual parameter value will fall.
- P.E ±(reliability factor × s.e)
- α= level of significance.
- 1- α= degree of confidence.

Desirable properties of an estimator

Unbiased
Expected value of estimator equals parameter e.g., E($\bar{x}$) = μ i.e, sampling error is zero.

Efficient
If var ($\bar{X}_1$) < var($\bar{X}_2$) then $\bar{X}_1$ is efficient than $\bar{X}_2$

Consistent
As n ↑, value of estimator approaches parameter & sample error approaches '0' e.g., As n → ∝ $\bar{x}$ → μ & s.e. → 0

Mean satisfies all 3 properties but median & mode don't.

If sample is not random then
- CLT does not apply.
- CI is not unbiased.
- Estimates will not have the desirable properties.

Student's T-Distribution
- Bell shaped.
- Shape is based on d.f.
- d.f. is based on 'n'.
- t-distribution depends on n (d.f) [normal distribution does not depend on n (d.f)].
- Symmetrical about it's mean.
- Less peaked than normal distribution.
- Has fatter tails.
- More probability in tails i.e., more observation are away from the centre of the distribution & more outliers.
- More difficult to reject H_0 using t distribution.
- C.I for a r.v. using t distribution must be wider when d.f are less.

Distribution		Variance		Sample		Test Statistic	
Normal	Non normal	Known	Unknown	Small (n<30)	Large (n≥30)	t	z
✓	×	✓	×	✓	×	×	✓
✓	×	✓	×	×	✓	×	✓
✓	×	×	✓	✓	×	✓	×
✓	×	×	✓	×	✓	✓*	×
×	✓	✓	×	✓	×	×	×
×	✓	✓	×	×	✓	×	✓
×	✓	×	✓	✓	×	×	×
×	✓	×	✓	×	✓	✓*	×

**The z-statistic is theoretically acceptable here, but use of the t-statistic is more conservative.*

Issues Regarding Selection of Appropriate Sample Size

As n ↑; s.e. ↓ & hence C.I is narrow.

Limitations of Large Sample Size

- Large sample may include observations from different population, hence:
  - Precision may not increase.
  - Precision may even fall.
- Cost may increase more relative to an increase in precision.

Biases

Data Mining Bias

Statistical significance of the pattern is overestimated because the results were found through data mining.

Data Mining

Using the same data to find patterns until the one that 'works' is discovered.

Warning Signs of Data Mining

- Evidence of testing many different, mostly unreported variables.
- Lack of economic theory consistent with empirical results.

Sample Selection Bias

- Systematically excluding some data from analysis.
- It makes the sample non-random.

Survivorship Bias

- Most common form of sample selection bias.
- Excluding weak performances.
- Surviving sample is not random.

Look –ahead Bias

Using sample data that wasn't available on the test date.

Time-period Bias

Time period over which the data is gathered is either too short or too long.

Study Session # 3, Reading # 11

"Hypothesis Testing"

$\propto$ = Level of significance
t.s = Test statistics
t.v = Table Value

t.v provides the critical values called as rejection points

s.s = Sample statistic
c.v = Critical value
s.e = Standard error

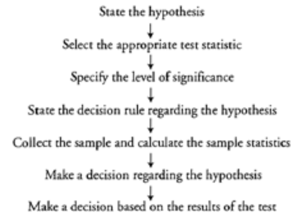
Hypothesis
Statement about parameter value developed for testing.

Two Types

Null Hypothesis H_0
- Tested for possible rejection.
- Always includes '=' sign.

Alternative Hypothesis H_a
The one that we want to prove.

Hypothesis Testing Procedure
- It is based on sample statistics & probability theory.
- It is used to determine whether a hypothesis is a reasonable statement or not.



(*Source:* Wayne W. Daniel and James C. Terrell, *Business Statistics, Basic Concepts and Methodology,* Houghton Mifflin, Boston, 1997.)

One Tailed Test
Alternative hypothesis having one side.
- Upper Tail
  $H_0: \mu \leq \mu_0$ vs $H_a: \mu > \mu_0$.
- Decision rule
  Reject H_0 if t.s > t.v.
- Lower Tail
  $H_0: \mu \geq \mu_0$ vs $H_a: \mu < \mu_0$.
- Decision rule
  Reject H_0 if t.s < - t.v.

Two Tailed Test
- Alternative hypothesis having two sides.
- $H_0: \mu = \mu_0$ vs $H_a\ \mu \neq \mu_0$.
- Reject H_0 if $|t.s| > t.v.$

Test Statistics (t.s)
Hypothesis testing involves two statistics:
- t.s calculated from sample data.
- critical values of t.s.

t.s is a random variable that follows some distribution.

Two Types of Errors

Type I Error
Rejecting a true null hypothesis.

Type II Error
Failing to reject a false null hypothesis.

Decision Rule
- It is based on distribution of t.s.
- It is specific & quantitative.

Significance Level (α)
- Probability of making a type I error.
- Denoted by Greek letter alpha (α).
- Used to identify critical values.

Power of a Test
- P(type II error).
- Probability of correctly rejecting a false null hypothesis.

p- value
- Probability of obtaining a critical value that would lead to a rejection of a true null hypothesis.
- Reject H_0 if p-value < α.

Relationship b/w Confidence Intervals & Hypothesis Tests
- Related because of critical value.

C.I
- [(s.s)- (c.v)(s.e)] $\leq$ parameter $\leq$ [(s.s) + (c.v)(s.e)].
- It gives the range within which parameter value is believed to lie given a level of confidence.

Hypothesis Test
- -c.v $\leq$ t.s $\leq$ + c.v.
- range within which we fail to reject null hypothesis of two tailed test given level of significance.

Statistical Significance vs Economical Significance
- Statistically significant results may not necessarily be economically significant.
- A very large sample size may result in highly statistically significant results that may be quite small in absolute terms.

σ^2 = population variance
N.dist = Normally distributed
N.N.dist = Non Normally distributed

n = sample size
$n \geq 30$ = large sample
n< 30 = small sample

t.s. = Test statistics
t.v = table value
d.f = degree of freedom

Testing	Conditions	Test Statistics	Decision Rule
Population Mean	• σ^2 known • N. dist.	$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$	• $H_0: \mu \leq \mu_0$ vs $H_a: \mu > \mu_0$ Reject H_0 if t.s. > t.v • $H_0: \mu \geq \mu_0$ vs $H_a: \mu < \mu_0$ Reject H_0 if t.s. < - t.v • $H_0: \mu = \mu_0$ vs $H_a: \mu \neq \mu_0$ Reject H_0 if | t.s | > t.v
	• $n \geq 30$ • σ^2 unknown	$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$ or $t^*_{n-1} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$ *(more conservative)	
	• σ^2 unknown • n<30 • N. dist.	$t_{n-1} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$; d.f = n-1	
Equality of the Means of Two Normally Distributed Populations based on Independent Samples.	Unknown variances assumed equal.	$t_{(n_1+n_2-2)} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ where; $s_P = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$ d.f = n_1+n_2 - 2	• $H_0: \mu_1 - \mu_2 \leq 0$ vs Ha: $\mu_1 - \mu_2 > 0$ Reject H_0 if t.s > t.v • $H_0: \mu_1 - \mu_2 \geq 0$ vs Ha: $\mu_1 - \mu_2 < 0$ Reject H_0 if t.s < -t.v • $H_0: \mu_1 - \mu_2 = 0$ vs Ha: $\mu_1 - \mu_2 \neq 0$ Reject H_0 if | t.s | > t.v
	Unequal unknown variances.	$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ $d.f = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2}}$	

Paired Comparisons Test

T.S $t_{(n-1)} = \frac{\bar{d} - \mu_{d0}}{s_{\bar{d}}}$

$\bar{d} = \frac{1}{n} . \sum d$

$s_{\bar{d}} = \frac{s_d}{\sqrt{n}}$

$s_d = \sqrt{\frac{\sum (d - \bar{d})^2}{n - 1}}$

Decision Rule

- $H_0: \mu_d \leq \mu_{d0}$ vs $H_a: \mu_d > \mu_{d0}$
  Reject H_0 if t.s > t.v.
- $H_0: \mu_d \geq \mu_{d0}$ vs $Ha: \mu_d < \mu_{d0}$
  Reject H_0 if t.s <-t.v.
- $H_0: \mu_d = \mu_{d0}$ vs $H_a: \mu d \neq \mu_{d0}$
  Reject H_0 if t.s > t.v.

Testing Variance of a N.dist. Population

T.s

$\chi^2_{(n-1)} = \frac{(n-1)s^2}{\sigma_0^2}$

Decision Rule

Reject H_0 if t.s > t.v

Chi-Square Distribution

- Asymmetrical.
- Bounded from below by zero.
- Chi-Square values can never be –ve.

Testing Equality of Two Variances from N.dist. Population

T.s

$F = \frac{s_1^2}{s_2^2}$; $s_1^2 > s_2^2$

Decision Rule

Reject H_0 if t.s > t.v

F- Distribution

- Right skewed.
- Bounded by zero.

Parametric Test

- Specific to population parameter.
- Relies on assumptions regarding the distribution of the population.

Non-Parametric Test

- Don't consider a particular population parameter.

Or

- Have few assumptions regarding population.

Study Session # 3, Reading # 12

"TECHNICAL ANALYSIS"

T.A = Technical Analysis
F.S = Financial Statements
C.F = Cash Flows
ROC = Rate of Change
RSI = Relative Strength Index

S.D = Standard Deviation
M.A = Moving Average
M.V = Market Value
MACD = Moving Average Convergence/divergence

12. a

- Study of collective market sentiment.
- Prices are determined by interaction of supply & demand.
- Key assumption of T.A is that EMH does not hold.
- Usefulness is limited in illiquid & outside manipulation markets.

Comparison	
Technical Analysis	Fundamental Analysis
• Share price & trading volume	• Intrinsic value
• Data is observable	• Use F.S & other information
• Can applied on assets without C.F	

12. b

- Charts of price & volume.
- Exponential price change ⇒ charts on a log scale.
- Time interval reflects horizon of interest.

Types of price charts

Line charts
- Show closing price as continuous line.

Bar charts
- Opening & high & low prices.
- Use dash & vertical lines as symbols.

Candlestick charts
- Same data as bar charts.
- Use box for opening & closing price.
- Patterns easier to recognize.

Point & figure charts
- Δ in direction of price.
- Horizontal axis no. of Δ in direction not time.
- Price increment chosen is "box size".

Relative strength analysis
- Assets closing price / Benchmark values.
- ↑ in trend, asset is outperforming & vice versa.

Volume Chart
- Usually displayed below price charts.
- Volume on vertical axis.

12. c

Trend in prices

Uptrend
- Reaching higher highs & retracing higher lows
- Show, demand is ↑
- Trend line connects increasing lows
- Breakout from downtrend (significant price change.

Downtrend
- Lower lows & retracing lower highs.
- Show, supply is ↑.
- Trend line connects decreasing highs.
- Breakdown from downtrend (significant price change).

Support Level
- Buying emerge.
- Prevent further price decline.

Change in polarity
- Breached resistance levels become support levels & vice versa.

Resistance Level
- Selling emerge.
- Prevent price rise.

Study Session # 3, Reading # 12

12. d

Reversal Patterns

- Trend approach a range of prices but fail to continue.
- Head & shoulder patterns for uptrend & inverse H&S for downtrends.
- Analyst use size of H&S pattern to project price target.
- Downtrend is projected to continue (price ↓ after the right shoulder forms).

Double top & triple top

- Indicate weakening the buying pressure. (Similar to H&S).
- Selling pressure appears after resistance level.
- Double bottom & triple bottom for down trends.

Continuation Patterns

Pause in trend rather reversal

Triangles

- Form when prices reach lower highs & higher lows.
- Can be symmetrical, ascending or descending.
- Suggest buying & selling pressure roughly equal.
- Size of triangle to set a price target.

Rectangles

- Form when trading temporarily form range b/w support & resistance level.
- Suggest prevailing trend will resume.
- Flags & pennants ⇒ short term price charts, rectangles & triangles.

12. e

Price based Indicators

Moving avg. lines

- Mean of last "n" closing prices.
- ↑in "n", smoother the avg. line.
- Uptrend price is higher then moving avg. & vice versa.
- M.A for different periods can be used together.
- Short term avg. above long term (golden cross) ⇒ buy signal, (dead cross) if vice versa ⇒ sell signal.

Bollinger bands

- Based on S.D of closing prices over last n periods.
- Analyst draw high & low bands above & below n-period M.A
- Prices above upper Bollinger ⇒ over bought, vice versa ⇒ over sold.
- Contrarian strategy ⇒ buy when most traders are selling.

Oscillators

- Tool to indentify overbought or oversold market.
- Based on M.P but scaled so they "oscillate" around a or between two values.
- Charts used to indentify convergence or divergence of oscillator & M.P
- Convergence ⇒ same pattern as price, divergence ⇒ vice versa.

Examples of oscillators

ROC or Momentum

- 100×Diff. b/w closing price & closing price n period earlier.
- Buy when oscillator from –to + & vice versa.
- Can be around o or around 100.

RSI

- Ratio of = $\frac{\text{Total price}\uparrow}{\text{Total price}\downarrow}$
- Oscillate b/w 0-100 .
- Value> 70 ⇒ overbought.
- Value < 30 oversold.

MACD

- Use exponentially smoothed M.V
- Oscillate around 0 but not bounded.
- MACD line above signal line ⇒ buy signal & vice versa.

Stochastic oscillator

- Calculated from latest closing price & highest & lowest prices.
- Use two lines bounded 0 & 100.
- %k = diff. b/w latest price & recent low as % of diff. b/w recent high & lows.
- % D line is 3-period avg. of the % k line.

12. e

Non-Price-Based Indicators

- Sentiment indicators ⇒ to gain insight into trends.
- Bullish⇒ increasing prices, bearish ⇒ decreasing prices.
- Opinion polls⇒ measure investor sentiment directly.

a. Sentiment Indicators

1. Put/Call Ratio

- Put/call ratio = $\frac{\text{Put volume}}{\text{Call volume}}$
- ↑Ratio ⇒ negative price outlook.
- Viewed as contrarian indicator.
- Extremely high ratio ⇒ bearish outlook & vice versa.

2. Volatility Index (VIX)

- Measure volatility of options on S&P 500 stock index.
- High VIX ⇒fear declines in stock market.
- Technical analysts interpret VIX in contrarian way.

3. Margin Debt

- ↑ in M.D,↑ buying, when reach their limit, buying↓, prices ↓, investor sell securities to meet margin calls.
- ↑ M.D coincides with ↑ prices & ↓ M.D with ↓ prices.

4. Short Interest Ratio

- Short interest is no. of shares borrowed & sold short.
- Short interest ratio = short interest / Avg. daily trading volume.
- ↑ Ratio, expect to ↓ in price & vice versa.

b. Flow of funds indicators

- Useful for observing changes in demand & supply of securities.

1. Arms index or short-term trading index (TRIN)

- Measure of funds flowing into advancing & declining stocks.
- $\text{TRIN} = \frac{\text{No.of adv.Issue /No.of declining issue}}{\text{Volume of adv.Issue/ volume of declining issue}}$
- Index value close to 1 ⇒ flowing evenly to advancing & declining stocks.
- Value > 1 ⇒ majority in declining stocks, value <1 ⇒ majority in advancing stocks.

2. Margin debt

- ↑ Margin debt ⇒ investor wants to buy more stocks & vice versa.

3. Mutual fund cash position

- Ratio of = fund's cash/Total assets.
- Uptrend ⇒ ratio ↓ & vice versa.
- T.A analysts view as contrarian indicator.

4. New equity issuance

- IPO & secondary issues add to supply of stocks.
- Issuer tends to issue when market peaks, so issuance coincide with high price.

12. f

- Cycle theory is study of processes that occur in cycles.

Cycle Periods

Presidential Cycles	Decennial Patterns	Kondratieff wave
4-year	10-year	18-year & 54 year

12. g

Elliott wave theory

- Financial market prices ⇒ described by interconnected set of cycles.
- Waves ⇒ chart patterns, uptrend ⇒ 5 upward waves, 3 down wards ↑ & vice versa in downtrend.
- Size of waves ⇒ Fibonacci ratios.
- Fibonacci numbers states with 0 & 1 and add each of previous two to produce next.
- Ratio of 0.618 and 1. 618 used to project price targets.

12. h

Intermarket analysis

- Analysis of interrelationship among M.V of asset classes (e.g. stocks, bonds)
- Relative strength ratio ⇒to identify outperformer asset class, then assets within class.