Correlation Analysis

$-1 \le r \le 1$

Scatter diagram $x\uparrow y\uparrow$ / $x\downarrow y\downarrow$ +ve $0 < r < 1$

Negative Correlation $x\uparrow y\downarrow$ / $x\downarrow y\uparrow$ $-1 < r < 0$

Perfect positive Correlation $x = y$, $r = 1$

Perfect Negative Correlation $r = -1$

No Correlation — Relation B/W intelligent and size of Shoe.

Karl Pearson Correlation Coefficient

If covariance is given

$$r = \frac{\text{cov}(x,y)}{\sigma_x \times \sigma_y}$$

If all obs are given

$$r = \frac{n\sum xy - \sum x \cdot \sum y}{\sqrt{N\sum x^2 - (\sum x)^2}\sqrt{N\sum y^2 - (\sum y)^2}}$$

$$\text{cov}(x,y) = \frac{\sum (x-\bar{x})(y-\bar{y})}{N}$$

Spearman Rank Correlation

$$r_R = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

$d_i = x_i - y_i$, n = no. of obs

If Ranks are Repeated

$$r_R = 1 - 6\left[\sum d_i^2 + \frac{1}{12}(m_1^3 - m_1) + \frac{1}{12}(m_2^3 - m_2)\right] \Big/ \; n^3 - n$$

$\boxed{\frac{1}{12}(m^3 - m)} \rightarrow$ correction factor

Coefficient of Concurrent deviations

$$r_{CCD} = \pm\sqrt{\pm\frac{(2C - m)}{m}}$$

C = No. of +ve deviation

$m = (n-1)$

Common property

$U = a + bx$ $\quad V = c + dy$

$r_{UV} = r_{xy}$ (sign of a) (sign of y)

$r_{vy} = r_{xy}$ (sign of a)

Probable Error

$$PE = 0.6745 \times \frac{1-r^2}{\sqrt{n}}$$

Standard Error

$$SE = \frac{1-r^2}{\sqrt{n}}$$

r = coefficient of Correlation, n = No. of obs

Properties of Probable Error

$|r| < 6PE \Rightarrow r$ is not significant

$|r| > 6PE \Rightarrow r$ is significant

Correlation of population $= r \pm PE$

Correlation of determination

Square of coefficient of Correlation r^2

$$r^2 = \frac{\text{Explained variance}}{\text{Total variance}} \rightarrow 1 - \frac{\text{unexplained}}{\text{Total variance}}$$

Coefficient of Non-determination $1 - r^2$

$$\frac{\text{Unexplained variance}}{\text{Total variance}} \rightarrow 1 - \frac{\text{explained}}{\text{Total variance}}$$

$\Rightarrow$ Both Coefficient of Determination r^2 & Coefficient of Non-determination $1 - r^2$ Can Never be -ve and can never exceed 1

Coefficient of Alienation

$$\sqrt{1 - r^2}$$

Regression Analysis

Aim 1 - Regression coefficient

Aim 2 - Regression Equation

Aim 3 - Estimating value of variable

calculate 1st term / calculate 2nd term

$\boxed{Y \text{ depends on } X}$ $\quad$ $\boxed{X \text{ depends on } Y}$

$Y = a + bx$ $\quad$ $X = a + bY$

$[Y - \bar{Y}] = b_{yx}[x - \bar{x}]$ $\quad$ $[X - \bar{X}] = b_{xy}[Y - \bar{Y}]$

If All obs are given

$\boxed{b_{yx}}$

$$b_{yx} = \frac{N\sum XY - \sum x\sum y}{N\sum x^2 - (\sum x)^2}$$

$\boxed{b_{xy}}$

$$b_{xy} = \frac{N\sum xy - \sum x\sum y}{N\sum y^2 - (\sum y)^2}$$

If r is and S.D are given

$$b_{yx} = r \cdot \frac{\sigma_y}{\sigma_x} \qquad b_{xy} = r \cdot \frac{\sigma_x}{\sigma_y}$$

If $\text{Cov}(x,y)$ is given

$$b_{yx} = \frac{\text{Cov}(x,y)}{\sigma_x^2} \qquad b_{xy} = \frac{\text{Cov}(x,y)}{\sigma_y^2}$$

$$\text{Cov}(x,y) = \frac{\sum(x-\bar{x})(y-\bar{y})}{N}$$

$$\sigma_y = \sqrt{\frac{\sum(y-\bar{y})^2}{N}}$$

$$\sigma_x = \sqrt{\frac{\sum(x-\bar{x})^2}{N}}$$

Common Property

Coefficient of Regression remains unchange due to change of origin but changes due to change of Scale

$$b_{uv} = b_{xy}\frac{N_x}{M_y} \qquad U = a + bx$$

$$b_{vu} = b_{yx}\frac{M_y}{N_x} \qquad V = c + dy$$

Make U and V Single

r is the GM of Regression coeffi

$$r = \sqrt{b_{yx} \times b_{xy}}$$

$$b_{yx} = r\frac{\sigma_y}{\sigma_x} \qquad b_{xy} = r\frac{\sigma_x}{\sigma_y}$$

$$\text{Correct } b_{xy} = \frac{1}{\text{incorrect } b_{yx}}$$

$$\text{Correct } b_{yx} = \frac{1}{\text{incorrect } b_{xy}}$$

$Y = a + bX$

a → Y-intercept; b → Regression Coefficient (slope)

Theoretical Distribution

Binomial distribution

$$P(x) = {}^{n}C_{x}(p)^{x}(q)^{n-x}$$

Mean = np

Variance (σ^2) = npq

S.D = $\sqrt{npq}$

Mode

Bi-Modal — $(n+1)p$ is integer

Mode: $(n+1)p$, $[(n+1)p]-1$

Uni-Modal — $(n+1)p$ is non integer

Mode = Largest value contained in $(n+1)p$

Poisson distribution

$$P(x) = \frac{e^{-m} \cdot m^{x}}{x!}$$

Normal distribution

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$x = \mu$ x ranges from $-\infty$ to ∞

Point of inflexion

$\mu+\sigma$ $\mu-\sigma$

QD : MD : SD   10 : 12 : 15

$Q_1 = \mu - 0.675\sigma$ $Q.D = \frac{Q_3-Q_1}{2}$

$Q_3 = \mu + 0.675\sigma$

Mean = $\mu_1 + \mu_2$

S.D = $\sqrt{\sigma_1^2 + \sigma_2^2}$

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

Index Numbers

Simple Avg $= \frac{\Sigma P_1}{\Sigma P_0} \times 100$

AM $= \frac{\Sigma P}{N}$ $\boxed{P = \frac{P_1}{P_0} \times 100}$

GM = Antilog $\left[\frac{\log P}{N}\right]$

Laspeyres

$$P_{01} = \frac{\Sigma P_1 q_0}{\Sigma P_0 q_0} \times 100$$

Paasche

$$P_{01} = \frac{\Sigma P_1 q_1}{\Sigma P_0 q_1} \times 100$$

Fishers

$$P_{01} = \sqrt{\frac{\Sigma P_1 q_0}{\Sigma P_0 q_0} \times \frac{\Sigma P_1 q_1}{\Sigma P_0 q_1}} \times 100$$

$\sqrt{L \times P}$

Dorbish & Bowley's Method

$$\frac{L+P}{2} \cdot \frac{\frac{\Sigma P_1 q_0}{\Sigma P_0 q_0} + \frac{\Sigma P_1 q_1}{\Sigma P_0 q_1}}{2} \times 100$$

Marshall Edgeworth

Avg of Qty $\frac{q_1+q_0}{2}$.

$$P_{01} = \frac{\Sigma P_1 (q_1+q_0)}{\Sigma P_0 (q_1+q_0)} \times 100$$

Value index

$V_{01} = \frac{V_1}{V_0} \times 100$ $V_{01} = P_{01} \times q_{01}$

Chain index Number

$$C.B.I = \frac{L.R\,CY \times C.B.I\ PY}{100}$$

Link relative $= \frac{CY\,P \times 100}{PY\ price}$ $\frac{P_1}{P_0} \times 100$

Conversion

FBI to CBI

$$C.B.I = \frac{FBI\ (CY)}{FBI\ (PY)} \times 100$$

CBI to FBI

$$FBI = \frac{C.B.I\ CY \times FBI\ PY}{100}$$

Link relative to price relative

$$\text{Price relative (CY)} = \frac{L.R\,(CY) \times \text{Price relative (PY)}}{100}$$

Shifting of Bases

New index No. Using New Base $= \frac{\text{Old Index Number using old base}}{\text{I.N of the year of New Base}} \times 100$

Splicing

Index no of old series under Forward Splicing

$$\frac{\text{Given IN of old series}}{\text{old series}} \times 100$$

Index no. of new series under Backward Splicing

$$\frac{\text{Given IN of New series} \times \text{Overlapping of old series}}{100}$$

Deflating

Real wage $= \frac{\text{Nominal wage}}{\text{Price index}} \times 100$

Money wage index $= \frac{\text{Real wage}}{\text{wages of Base year}} \times 100$

Real wage index $= \frac{\text{Money wage}}{\text{Price index}} \times 100$

$$\frac{\text{Current value}}{\text{Price index}} \times 100$$

Consumer Price Index

Aggregate Expenditure Method $= \frac{\Sigma P_1 q_0}{\Sigma P_0 q_0} \times 100$

Family Budget Method $= \frac{\Sigma PW}{\Sigma W}$

$P = \frac{P_1}{P_0} \times 100$ W = weight

Given = $\sqrt{}$ Not given = $P_0 q_0$

Industrial price index $= \frac{\Sigma IW}{\Sigma W}$

Purchase power $= \frac{1}{P.I}$ P.I = price index

Types of Test

1. Unit Test. Formula of index number should be free of units
2. Time Reversal Test $P_{01} \times P_{10} = 1$
3. Factor Reversal Test $P_{01} \times q_{01} = V_{01}$
4. Circular Reversal Test $P_{01} \times P_{12} \times P_{20} = 1$

FORMULA SHEET

Measures of central Tendency

Mean ($\bar{x}$)

individual data - $\bar{x} = \frac{\sum x_i}{N}$

discrete data $\bar{x} = \frac{\sum f_i x_i}{N}$ $N = \sum f_i$

Continuous data $\bar{x} = \frac{\sum f_i x_i}{N}$ x_i = Mid point

Correct Mean

$$\bar{x} = \frac{\sum x_i \text{ incorrect} - \sum \text{incorrect} + \text{Correct}}{N}$$

Combined Mean

$G_1 = n_1\bar{x}$ $G_2 = n_2\bar{x}$

$$\bar{x}_{12} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1+n_2}$$

Individual series

Median = $\left(\frac{n+1}{2}\right)^{th}$ observation

$Q_K = K\left(\frac{n+1}{4}\right)^{th}$ obs $D_K = K\left(\frac{n+1}{10}\right)^{th}$ obs

$P_K = K\left(\frac{n+1}{100}\right)^{th}$ obs

Discrete series - same as above

Continuous series

Median = $L + \left[\frac{\frac{KN}{2} - CF_p}{f}\right] \times i$

$Q = \frac{KN}{4}$ $D = \frac{KN}{10}$ $P = \frac{KN}{100}$

Mode - Value repeats Maximum no. of times

Individual - Maximum, no. of times

Discrete - highest frequency

Continuous - $L + \frac{f_m - f_1}{2f_m - f_2 - f_1} \times i$

GM individual = $(x_1 \times x_2 \times x_3 \cdots)^{\frac{1}{n}}$

Discrete = Antilog$\left[\frac{\sum f_i \log x_i}{N}\right]$

How to find Log?

$\sqrt{a}$ - 1 × 3558 =
13 times

Antilog

$\sqrt{10}$ - 1 × Antilog + 1 × =
12 times 12 times

GM of z = GM of x × GM of y

GM of z = $\frac{\text{GM of } x}{\text{GM of } y}$

HM individual data = $\frac{N}{\sum \frac{1}{x_i}}$ $N = \sum f_i$

discrete data = $\frac{N}{\sum \frac{f_i}{x_i}}$

HM Between two numbers a and b is

$$HM = \frac{2ab}{a+b}$$

→ Average of two ratio eg. Speed = $\frac{\text{Distance}}{\text{Time}}$

Combined HM

G_1 G_2

n_1 n_2

H_1 H_2

$$HM = \frac{n_1+n_2}{\frac{n_1}{H_1}+\frac{n_2}{H_2}}$$

Relationship B/w Mean, Median, Mode

Symmetrical data = Mean = Median = Mode

Asymmetrical data
- Positively skewed: Mean > Median > Mode
- Negatively skewed: Mean < Median < Mode

Moderately skewed

Mode = 3Median - 2 Mean

Relationship B/w AM GM HM

Symmetrical AM = GM = HM

Asymmetrical AM > GM > HM

if Qs. did not give anything we inequality AM ≥ GM ≥ HM

Relations B/w AM GM HM

$GM^2 = AM \times HM$

Common property

MOCT gets affected due to change of origin as well as scale

$y = a + bx$ property

origin scale $Y_{MOCT} = a \times b X_{MOCT}$

MOCT
- Mean
- Median
- Mode
- GM/HM

Measures of Dispersion

Range = L-S

Coefficient of Range = $\frac{L-S}{L+S} \times 100$

Mean Deviation

$MD_A = \frac{\sum |x - A|}{N}$ A = Mean / Median / Mode

discrete Series

$MD_A = \frac{\sum f_i |x_i - A|}{N}$ $N = \sum f_i$

Coefficient of Mean deviation $= \frac{MD_A}{A} \times 100$

Quartile deviation $QD = \frac{Q_3 - Q_1}{2}$

Coefficient of QD = $\frac{Q_3 - Q_1}{Q_3 + Q_1}$

Coefficient of Quartile Variation = $\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$

Standard Deviation

$\sigma = \sqrt{\frac{\sum x^2}{N} - (\bar{x})^2}$ direct Method

$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{N}}$ (indirect Method)

Coefficient of Variance $= \frac{\sigma}{\bar{x}} \times 100$

discrete series $\sigma = \sqrt{\frac{\sum f_i (x - \bar{x})^2}{N}}$

Continuous same as discrete by taking Mid point

Combined standard deviation

$$\sigma_{12} = \sqrt{\frac{n_1\sigma_1^2 + n_2\sigma_2^2 + n_1 d_1^2 + n_2 d_2^2}{n_1+n_2}}$$

$d_1 = \bar{x}_{12} - \bar{x}_1$ $d_2 = \bar{x}_{12} - \bar{x}_2$

Variance $(S.D)^2 = \sigma^2$

Relationship B/w MD, QD, SD

QD : MD : SD = 10:12:15

Mathematics of Finance

Compound interest

$C.I = \frac{R}{100} + 1 \times P = A - P = Int$ (↓ No. of Years)

Yearly, Half Yearly, Quarterly, Monthly

Amt in C.I forms G.P

$P, A_1, A_2, A_3, \ldots = GP$

$r - 1 = i$

$A_2, A_4, A_6, A_8 \ldots = G.P$

$(r)^{\frac{1}{n}} - 1 = i$

Forward Term of GP $\times r$

Backward Term of GP $\div r$

Diff B/W C.I & S.I

$CI_2 - SI_2 = PR^2$

$CI_3 - SI_3 = PR^2[R+3]$

$CI_n - SI_n = P\left[\{(1+1)^n - 1\} - in\right]$

Independent of Principal

$R = 2\left[\frac{CI_2 - SI_2}{SI_2}\right]$

Time requires to double an investment

$T = 0.35 + \frac{0.69}{R\%}$

Rates changing every year

$P + R_1\% + R_2\% + R_3\% \ldots = A_n$

$A_q = (m)^{\frac{q}{n}}$ Sum of money amt to it will amt to ?

"m" times "q" years

"n" year

Depreciation

$[1 - r\%] \times MV = Scrap$ (↓ Life of Machine)

Present Value

Ordinary annuity

A to PV: $\frac{R}{100} + 1 \div = GT \times a = PV$ (↓ No. of installment)

PV to Annuity: $\frac{R}{100} + 1 \div = GT\,(m+)$ (↓ No. of install) $\frac{PV}{MRC} = a$

Annuity due

A to PV: $1 + = \frac{R}{100} + 1 \div = GT \times a = PV$ (↓ $(n-1)$)

PV to Annuity: $1 + = \frac{R}{100} + 1 \div = GT\,m+$ (↓ $(n-1)$) $\frac{PV}{MRC} = a$

Future Value

Ordinary Annuity

a to FV: $1 + = \frac{R}{100} + 1 \times 1 = GT \times a = FV$ (↓ $(n-1)$)

FV to a: $1 + = \frac{R}{100} + 1 \times 1 = GT\,m+$ (↓ $(n-1)$) $\frac{FV}{MRC} = a$

Annuity due

A to FV: $\frac{R}{100} + 1 \times 1 = GT \times a = FV$ (↓ No. of installment)

FV to a: $\frac{R}{100} + 1 \times 1 = GT\,(m+)$ (↓ no. of install) $\frac{FV}{MRC} = a$

Yearly, Halfyearly, Quarterly, Monthly

Perpetual Annuity

	ordinary	Annuity Due
Without growth	$PV_\infty = \frac{a}{i\%}$	$PV_\infty = a + \frac{a}{i\%}$
With growth	$PV_\infty = \frac{a}{i-g}$	$PV_\infty = a + \frac{a}{i-g}$

Annually	Half Yearly	Quarterly	Monthly
$r\%$	$\frac{r\%}{2}$	$r\%/4$	$\frac{r\%}{12}$

Applications of TVM

PV
1. Net present value
2. Capital investment
3. Leasing Concept
4. Valuation of Bond

FV
1. Sinking fund

Net Present value = PVCI − PVCO (income) − (Expense)

Same cash inflow — A to PV trick: $\frac{R}{100} + 1 \div = GT \times a = PV$ (↓ No.)

different cash inflow → Table concept

Present Value of Future Money (principal) = Discounting factor of that year × FM

Value of Bond = P.V of int + P.V of Redemption value

$[a \to PV]$ $[\text{discounting factor} \times RV]$

Nominal Rate of Return

$NRR = RRR - [-\text{inflation}]$

$= RRR + \text{inflation}$

Real Rate of Return

$RRR = NRR + [-\text{inflation}]$

$= NRR - \text{inflation}$

CAGR $= \left[\left[\frac{CV}{IV}\right]^{\frac{1}{\text{no. of period}}} - 1\right] \times 100$

CV = Current value

IV = Initial Value.

$(n-1)$

Amount $= P[1+i]^n$

$C.I = P\left[[1+i]^n - 1\right]$

Effective Rate of interest $= (1+i)^n - 1$

$PV = \frac{a}{i}\left[1 - \frac{1}{(1+i)^n}\right]$

$(1+i)^n \to$ discounting factor

$FV = a\left[\frac{(1+i)^n - 1}{i}\right]$

Annuity due > Ordinary Annuity

Sequence & Series

AP → $n^{th} = a + (n-1)d$

$Sum = \frac{n}{2}[2a + (n-1)d]$

$S_n = \frac{n}{2}[a + l]$

GP → $n^{th} = ar^{n-1}$

$Sum = \frac{a(1-r^n)}{1-r}$ $1 > r$

$= \frac{a(r^n - 1)}{r - 1}$ $r > 1$

$S_\infty = \frac{a}{1-r}$ $|r| < 1$