Correlation

Dr. Abhijit Ghosh
Assistant Professor of Economics
A N Sinha Institute of Social Studies

To be covered

- Definition
- Correlation coefficients
- Example

Correlational Research

- Definition
  - Whether and to what degree variables are related
- Purpose
  - Determine relationships
  - Make predictions
- Limitation
  - Cannot indicate cause and effect

The correlation coefficient between X and Y denoted by r is
$r = cov(X,Y) / \sigma x \sigma Y$

σx σY are the standard deviation of X and Y

Cov (X, Y) denotes the covariance of X and Y
This is also known as Pearson correlation.

Properties

1. Correlation Coefficient (r) is independent of origin and scale of the observations.

Let x and y be the variables and (x1, y1), (x2, y2), (x3, y3) ..... (xn, yn) are the n pairs of observation. Let us change the of x and y to a and c respectively and units of measurement (i.e. scale) to b and d respectively Then ui=(xi-a)/b vi= (yi-c)/d

Where a, b, c, d are arbitrary constants (b and d positive)
Then rxy= ruv

2. r is a pure number.

3. r lies between -1 and +1

Correlation Coefficients

- Interpreting the size of correlations
  - General rule
    - Less than .35 is a low correlation
    - Between .36 and .65 is a moderate correlation
    - Above .66 is a high correlation

Statistical Significance

- Sample size and statistical significance
  - Small samples require higher correlations for significance
  - Large samples require lower correlations for significance

Types of Correlation Coefficients

- The type of correlation coefficient depends on the measurement level of the variables
  - Pearson *r* - continuous predictor and criterion variables
    - Math attitude and math achievement
  - Spearman rho – ranked or ordinal predictor and criterion variables
    - Rank in class and rank on a final exam

Differences between Types of Studies

- Correlational research is a general category that is usually discussed in terms of two variables
- Relationship studies develop insight into the relationships between several variables
  - The measurement of all variables occurs at about the same time
- Predictive studies involve the predictive relationships between or among variables
  - The predictor variables are collected long before the criterion variable

Variables

There should be rationality in selecting Variables to be correlated.

Correlation Coefficients

correlation coefficient lies between 1 to − 1

- Direction
  - Positive or negative

Scatter Diagram: A scatter Diagram shows the nature of association between the two variables.

Correlation Coefficient (r) : Let (x1, Y1), (X2, Y2), (X3, Y3)..... (xn, yn) be a given set of n pairs of observation on two variables X and Y

Variables

There should be rationality in selecting Variables to be correlated.

Scatter Diagram: A scatter Diagram shows the nature of association between the two variables.

Correlation Coefficient (r) : Let (x1, Y1), (X2, Y2), (X3, Y3)..... (xn, yn) be a given set of n pairs of observation on two variables X and Y

The correlation coefficient between X and Y denoted by r is
$r = cov(X,Y) / \sigma x \sigma Y$

σx σY are the standard deviation of X and Y

Cov (X, Y) denotes the covariance of X and Y
This is also known as Pearson correlation.

Properties

1. Correlation Coefficient (r) is independent of origin and scale of the observations.

Let x and y be the variables and (x1, y1), (x2, y2), (x3, y3) ..... (xn, yn) are the n pairs of observation. Let us change the of x and y to a and c respectively and units of measurement (i.e. scale) to b and d respectively Then ui=(xi-a)/b vi= (yi-c)/d

Where a, b, c, d are arbitrary constants (b and d positive)
Then rxy= ruv

2. r is a pure number.

3. r lies between -1 and +1.

Rank Correlation (R)

Situation may arise where precise measurements are not available or characters cannot be measured at all.

Suppose, the marks obtained in Mathematics and Chemistry by 30 students of Standard XII are given. Now the question is whether marks obtained in two subjects related. This can be calculated using Spearmen rank correlation.
$R = 1 - [6\Sigma d2 / (n3 - n)]$

D is the difference of the ranks of the individual student in two subjects (as in example);
n= total number of students
R lies between -1 and +1.

Thank You

Correlation

Abhijit Ghosh

[Read any standard Statistics book, all book must cover this portion, feel free to communicate any problem]

Correlation. Correlation is a **statistical** technique that shows whether and how strongly pairs of variables are related.

For example, height and weight are related;

Scatter Diagram: A scatter Diagram shows the nature of association between the two variables.

Correlation Coefficient (r) : Let (x1, Y1), (X2, Y2), (X3, Y3)..... (xn, yn) be a given set of n pairs of observation on two variables X and Y

The correlation coefficient between X and Y denoted by r is

$$r = \frac{cov\,(X,Y)}{\sigma x \sigma Y}$$

σx σY are the standard deviation of X and Y

Cov (X, Y) denotes the covariance of X and Y

This is also known as Pearson correlation.

Properties

1. Correlation Coefficient (r) is independent of origin and scale of the observations.

   Let x and y be the variables and (x_1, y_1), (x_2, y_2), (x_3, y_3) ..... (x_n, y_n) are the n pairs of observation. Let us change the of x and y to a and c respectively and units of measurement (i.e. scale) to b and d respectively
   Then $u_i=(x_i-a)/b$ $v_i=(y_i-c)/d$
   Where a, b, c, d are arbitrary constants (b and d positive)
   Then $r_{xy}= r_{uv}$

2. r is a pure number.

3. r lies between -1 and +1.

Rank Correlation (R)

Situation may arise where precise measurements are not available or characters cannot be measured at all.

Suppose, the marks obtained in Mathematics and Chemistry by 30 students of Standard XII are given. Now the question is whether marks obtained in two subjects related. This can be calculated using Spearmen rank correlation.

$$R = 1 - [6\Sigma d^2 / (n^3 - n)]$$

D is the difference of the ranks of the individual student in two subjects (as in example);

n= total number of students

R lies between -1 and +1.

Assignment

1. What is the significance of correlation analysis in social science research?
2. Suppose in any teste you get r=2. How do you explain the result?
3. What is spurious (nonsense) correlation?
4. Calculate Pearson's correlation coefficient of the given data:

X	Y
65	68
63	66
67	68
64	65
68	69
62	66
70	68
66	65

Explain the result.

5. Prove that the correlation coefficient does not depend on the origin and the scale of the observation.

6. Calculate Rank Correlation

In a certain examination 10 students obtained the following marks in Mathematics and Physics. Find Spearman's rank correlation coefficient.

Student (Roll No.)	Marks in Mathematics	Marks in Physics
1	90	85
2	30	42
3	82	75
4	45	68
5	32	45
6	65	63
7	40	60
8	88	90
9	73	62
10	66	58

Dr. Abhijit Ghosh

Q Prove that the correlation coefficient (r) does not depend on the origin or scale of the observation.

Proof :- Let x and y be the variables and $(x_1, y_1), (x_2, y_2) \ldots\ldots (x_n, y_n)$ denote n pairs of observations. Let us change the origins of x and y to a and c, and the units of measurement (i.e. scales) to b and d respectively

So, $u_i = \frac{x_i - a}{b}$ $\quad v_i = \frac{y_i - c}{d}$ $\quad \ldots (i)$

where a, b, c, d are arbitrary constants (b and d positive). We have to prove that the correlation coefficient between x and y is the same as that between u and v; i.e. $r_{xy} = r_{uv}$

Scanned with CamScanner

From (i) $x_i = a + bu_i$; $y_i = c + dv_i$

Therefore $\bar{x} = \Sigma x_i / n$

$= \dfrac{\Sigma (a + bu_i)}{n}$

$= (na + b\Sigma u_i)/n$

$= a + b\bar{u}$

Similarly $\bar{y} = c + d\bar{v}$

Hence $\sigma_x^2 = \Sigma (x_i - \bar{x})^2 / n$

$= \Sigma (a + bu_i - a - b\bar{u})^2 / n$

$= b^2 \Sigma (u_i - \bar{u})^2 / n$

$= b^2 \sigma_u^2$

$\therefore \sigma_x = b\,\sigma_u$ ($\because$ b is positive)

$\sigma_y = d\,\sigma_v$ ($\because$ d is positive) } — (iii)

Scanned with CamScanner

Again $\text{Cov}(x,y) = \dfrac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{n}$

$$= \frac{\Sigma\{(a+bu_i - a - b\bar{u})\{c+dv_i - c - d\bar{v}\}\}}{n}$$

$$= bd\,\Sigma(u_i - \bar{u})(v_i - \bar{v})/n$$

$$= bd\ \text{cov}(u,v) \longrightarrow (iii)$$

Substituting the values from (ii) and (iii)

$$r_{xy} = \frac{\text{cov}(x,y)}{\sigma_x \sigma_y} = \frac{bd\ \text{cov}(u,v)}{(b\sigma_u)(d\,\sigma_v)}$$

$$= \frac{\text{cov}(u,v)}{\sigma_u \sigma_v}$$

$$= r_{uv}$$

$\therefore r_{xy} = r_u$ (Proved)

Scanned with CamScanner

Growth Perspectives

Dr. Abhijit Ghosh

Introduction

All economic models are built on certain assumptions, which can be derived from some primitives. However, it is necessary to understand the essential dynamics of a model in order to debate its contribution in the greater context of human learning. Media persons debate regularly about the problems and prospects of long-run growth. Politicians like to project themselves as pro-growth. Policy makers analyse impact of government decisions on the economic growth. Uneven growth may lead to social and political tensions. In fact, the huge amount of inequality generated in the present-day world is an indirect fall out of growth. Poverty, malnutrition and other deprivations become intolerable and unjustifiable in an atmosphere of sustained growth. One of the causes of the demise of East European communist regimes is their inability to sustain growth. Currently scientists are concerned about the possible impact of growth on environment and resource base of the world. Economists since the time of Adam Smith are concerned about growth. The main debate between him and his predecessors (mercantilists and physciocrats) was mainly based on the sources of economic prosperity. It so fascinated him that it was stamped on the cover of his famous treatise An Inquiry into the Nature and causes of the Wealth of Nations. However the prediction of a gloomy future by Malthus (1826) earned economics its name dismal science. Marx analysed the evolution of history as a conflict between growth sustaining efforts and that which wishes to strangle it. However modern growth theory started with Harrod (1939) and Domar (1946). This Keynesian perspective was replaced by the neoclassical view by Solow (1956) and Swan (1956). After a development for two decades the subject seemed to have met its natural death. It was the new growth theory that revitalized the field utilizing some earlier works (Arrow, 1962; Shell, 1967). However, the subject is couched in a mathematical form that is not easily accessible to ordinary people whose life are intimately linked with it. It is thus necessary to spell out the empirics of growth that helps us to gather the fire and excitement in studying growth that fascinated economists and philosophers since time immemorial." (source: Sengupta et. al 2015)

Table 1. Growth of per capita world GDP (in 1985 dollars)

Periods	Growth Rate	History
500-1500	0%	Medieval period
1500-1750	0.08%	Period of Merchant capital
1750-1850	0.17%	First phase of Industrial Revolution
1850-1950	0.88%	Second phase of Industrial Revolution
1950-1990	2.20%	Global spread of growth phenomenon

Source: Charles I. Jones: Introduction to Economic Growth

Classical Growth Theory: This theory covers a wide range of individual thinkers including Adam Smith, David Ricardo, James Mill and others. While they differ in many points they all argue land as a limiting factor of growth.

Keynesian School: This school emphasizes on the demand-supply adjustments in the growth process.

Neo-Classical Theory of Growth: The neoclassical school introduces capital as a factor that can be accumulated by human effort. However diminishing returns to capital finally frustrate the growth attempt.

New Growth Theory: The new growth school replaces capital by intermediate goods & human capital. Growth is a result of conscious human activity.

Knife-Edge Instability

Knife-edge equilibrium is an unstable equilibrium. If that equilibrium is reached, you may continue to stay in that equilibrium in the absence of any slightest shock or disturbance. Once such a small disturbance, moves the system away from the equilibrium, the system cannot ensure automatic return to the same equilibrium. Thus the underlying forces do not restore back to the original equilibrium by self-correcting process.

One of the most famous examples of the knife's edge equilibrium is in the Harrod-Domar Growth Model (1949), which sought to integrate some theory of economic growth with the Keynesian General Theory. Under the Harrod-Domar model, there is a precise rate of investment which is compatible with full employment.

Actual growth

Actual growth is the real **rate increase** in a country's GDP per year. **Natural growth** is the **growth** an economy requires to maintain full employment.

The warranted growth rate

The warranted growth rate is the growth rate at which all saving is absorbed into investment. If, for example, people save 10 percent of their income, and the economy's ratio of capital to output is four, the economy's warranted growth rate is 2.5 percent.

Inada conditions

In macroeconomics, the **Inada conditions**, named after Japanese economist Ken-Ichi Inada, are assumptions about the shape of a production function that guarantee the stability of an economic growth path in a neoclassical growth model.

It can be shown that the Inada conditions imply that the elasticity of substitution is asymptotically equal to one (although the production function is *not* necessarily asymptotically Cobb–Douglas).

In stochastic neoclassical growth model, if the production function does not satisfy the Inada condition at zero, any feasible path converges to zero with probability one provided that the shocks are sufficiently volatile.

Bibliography

- Aghion, P., & Howitt, P. (1992, March). A model of growth through creative destruction. Econometrica, 60(2), 323–351.
- Arrow, K. J. (1962). Economic welfare & the allocation of resources for invention. In The rate & direction of inventive activity: Economic & social factors (pp. 609-626). NBER.
- Barro, R. J., & Sala-i-Martin, X. (1997). Convergence. Journal of Political Economy, 100(2),
- Dixit, A. K., & Stiglitz, J. E. (1977). Monopolistic competition & optimum product diversity. The
- American Economic Review, 67, 297–308.
- Domar, E. (1946). Capital expansion, rate of growth, & employment. Econometrica, 14(2), 137–147.
- Grossman, G., & Helpman, E. (1991). Innovation & growth in the global economy. Cambridge, MA: MIT Press.
- Hansen, G. D., & Prescott, E. C. (1998). Malthus to Solow. NBER Working Papers 6858. National
- Bureau of Economic Research, Inc.
- Harrod, R. F. (1939). An essay in dynamic theory. The Economic Journal, 49(193), 14–33.
- Harrod, R. F. (1948). Towards a dynamic economics. London: Macmillan.
- Jones, C.I. (2001). Introduction to economic growth (2nd ed.). W. W. Norton & Company.
- Romer, P. M. (1990). Endogenous technological change. Journal of Political Economy, 98(S5), S71–S102
- Samuelson, P. A. (1948). Economics: An introductory analysis. McGraw-Hill.
- Swan, T. (1956). Economic growth & capital accumulation. The Economic Record, 32(2), 344–361.
- Sengupta, A and Anirban Hazra (2015): Growth Theory: A Primer in Roy, N et. Al. (ed.) Strategic Infrastructure Development for Economic Growth and Social Change, Business Science Reference

Human Development Paradigm

Dr. Abhijit Ghosh

A N Sinha Institute of Social Studies

Some Facts

COUNTRY	INCOME RANK	HDI RANK
India	133	136
Kyrgyzstan	149	124
Equatorial Guinea	39	136

Engaged Research

After 15 years of research on hunger and famines, one is perhaps entitled to feel like an 'expert' of sorts on these matters. Yet I did not always find myself better equipped than others to understand the practical issues...At times, I even felt embarrassingly ignorant compared with local people who had little formal education but a sharp understanding of the real world. (Drèze)

Engaged Research – virtuous circle for research quality

'the value of scientific research can, in many circumstances, be enhanced *even further* if it is combined with real world involvement and action.'

Jean Drèze

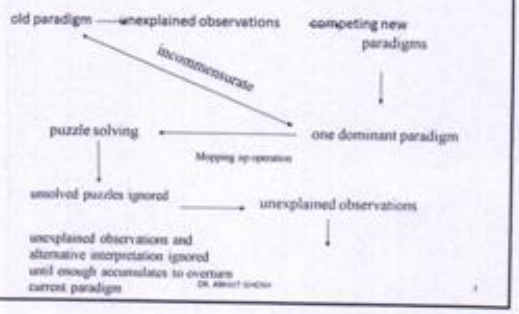
Human Development Paradigm

- Conceptual Issues
- Measurement Issues

Content

- Historical Evolution
- Note key antecedents – Basic Needs (BN) and Capabilities
- Human Development Report
- Importance in Social Science Research

Paradigm Diagram



Historical evolution of thought about development:post-colonialism

- Colonial legacy
  - low incomes
  - low savings/investment
  - appalling health; low literacy
  - primary production, with little manufacturing
  - total dependency

Development thinking on gaining independence

- Major emphasis on economic growth (Rostow; Lewis; Hirschman).
- Using 'surplus' labour. (Nurkse; Lewis)
- On industrialisation and raising investment as mechanisms. (Mahalanobis; Rosenstein-Rodan)
- In parallel, emphasis on costs of dependence (Prebish/Singer; Frank; Amin; Sunkel).
- General neglect of 'human' dimension – thought it would be looked after automatically through growth.

But challenge to growth followed

Why

- Rising unemployment; high underemployment
- Poverty still high and increasing in absolute terms
- Dependency remained – finance, technology, management.
- Growth of GNP neglects income distribution, public goods, employment – all essential for improving quality of life.

Defects of focus on money incomes.

- Income distribution critically important
- Neglects public goods (and externalities more generally).
- Assumes utilitarian philosophy. Money income measures utility. Utility should be maximised.
- 'Physical condition neglect'. Entrenched deprivation can become acceptable.
- Is consequentialist. Neglects agency goals (how you get there – e.g. child labour).
- Assumes people only consider OWN welfare.

Succession of challenges to growth only agenda from 1970s

- 'Dethronement' of GNP (Seers)
- Employment objective emphasised: ILO missions to Colombia, Sri Lanka – led by Seers.
- Redistribution with growth (Chenery, Singer and others).
- Basic needs (ILO, WB, Ghai, Streeten, Ranis, Stewart).
- Capabilities (Sen)
- Human Development

Explaining evolution

- Employment objective of ILO early 1970s: But why employment? Employment is a *means* to achieve various objectives, including incomes, production and recognition
- Hence move to focus on incomes of the poor. Redistribution with growth (RWG).
- Is the strategy feasible technically? Politically?
- Is it right to focus on money incomes?

DR. ABHIJIT GHOSH 11

Evolution of the concept of Human Welfare

- ❑Concept of Human Welfare in Upanishad (Sen, 1999)
- ❑Thought of Aristotle
- ❑Thought of Emmanuel Kant
- ❑Thought of Adam Smith
- ❑Thought of Bentham
- ❑Thought of Marshall
- ❑PQLI Approach(Dobb 1979)
- ❑Basic Needs Approach
- ❑Human Development

DR. ABHIJIT GHOSH 12

Evaluation of human condition

- Mental satisfaction (Utilitarian approach)
- Command on basic Commodities (Rawls)
- Entitlement approach (Sen)
- Extended Entitlement (Feminist economics)
- Capability Approach

DR. ABHIJIT GHOSH 13

"Entitlement is a legal concept, dealings with rules that govern who can have the use of what" (Sen 2004).

"Extended entitlement" as defined by Sen broadens the concept of entitlement by invoking the concept of "*perceived legitimacy*" rather than only the law in the strict sense.

Basic Need Approach

Provisions of health services, education, housing, sanitation, water supply and adequate nutrition

Physical Quality of Life Index

Mortality, Life Expectancy, Basic Literacy

Capability Approach

- Sen's capability approach is a moral framework. It suggests that **social arrangements should be primarily evaluated according to the extent of freedom people have to promote or achieve functionings they value.**

An Example (Dasgupta 2008)

- Story of two girls
- Desta in rural Ethiopia→ destined to be married to a poor farmer like her father→ no mobility
- Becky in suburban town of American Midwest→ destined to be a doctor -> move up in the mobility ladder
- Two "intrinsically" similar girls face two different positional objectivity

Capability

- **the various combinations of functionings (beings and doings) that the person can achieve. [It] is, thus, a set of vectors of functionings, reflecting the person's freedom to lead one type of life or another...to choose from possible livings.** *(Inequality Re-examined)*
- *think of it as a budget set*

All formulations of capability have two parts: freedom and valuable beings and doings (functionings). Sen's significant contribution has been to unite the two concepts.

Functionings

the various things a person may value and have reason to value doing or being

- intuitive
- intrinsically valuable to the person
- intrisic value (have reason to value)
- so avoids adaptive preferences
- 'doings and beings' is our focal *space*

Functionings allow for different interpersonal conversion factors

Resources → Capability → Functionings → Utility

Bike	Able to ride around	Ride around	☺
Food	Able to be nourished	Nourished	☺

Note: functionings & indicators

- Which are direct indicators of functionings?
  A. Asset index
  B. Subjective Well-being / Happiness
  C. Body Mass Index
  D. Literacy
  E. Years of Schooling
  F. Self-reported health

Freedom

- "the *real opportunity* that we have to accomplish what we value"
- "The 'good life' is partly a life of genuine choice, and not one in which the person is forced into a particular life – however rich it might be in other respects."

It is *authentic self-direction* – the ability to shape one's own destiny as a person and a part of various communities.

Sen-Haque Debate (Fukuda-Parr and Shiva Kumar, 2004)

- Haque insisted a unidimension index that can capture the extent of human deprivation.
- Sen argues that such an index is vulgar.
- A unidimensional deprivation index is less vulgar than GDP
- Sen argues that Human development is a very complex issue that can not be aggregated.
- Officials believe more in indices than in the plathora of tables.
- Sen admits a unidimensional index is practical but he argues, HDI has to move beyond it.

HDI: Basic Framework

- A composite index consisting of three indices:
  - Gross Domestic Product (GDP) Index
  - Life expectancy (LE) Index
  - Education (E) Index

HDI: Basic Framework

- All indices are normalized between zero and one
- GDP Index=

$$\frac{\log(\text{PCGDP}) - \log(\$100)}{\log(\$40{,}000) - \log(\$100)}$$

- Life Expectancy Index= $\frac{LE - 25}{85-25}$

HDI: Basic Framework

- Education Index
  - Consists of two sub-indices
- Adult literacy (AL) index
- Gross school enrolment (GSE) index

Adult literacy index =

$$\frac{AL-0}{100-0}$$

- Gross school enrolment index =

$$\frac{GSE - 0}{100-0}$$

Education Index = 2/3 Adult literacy index + 1/3 Gross school enrolment index

HDI: Basic Framework

HDI =1/3 GDP Index +1/3 Life exp Index +1/3 Education Index

Example

- Per Capita Gross Domestic Product = $3, 139
- Life expectancy = 63.6 years
- Adult literacy = 61%
- Gross school enrolment = 62%
- Per Capita gross domestic product (PGDP) index =

$$\frac{\log(\$3,139) - \log(\$100)}{\log(\$40,000) - \log(\$100)} = 0.58$$

- Life expectancy = (63.6 25) / (85 25) = 0.64
- Adult literacy index = (61 0) / (100 0) = 0.61
- Gross school enrolment index = (62-0)/(100 - 0) = 0.62
- Education index = 2/3 0.61 + 1/3 0.62 = 0.61

HDI = 1/3 0.58 + 1/3 0.64 + 1/3 0.61 = 0.61

HDI: Criticism

- HDI is not sensitive to inequality across persons
- Linear Aggreagation
-

Some key aspects of HD

1. Humans are ends not means
2. In practice major focus is on BN type goods and services, but also discusses other issues (freedom, democracy, gender, environment, communities, culture) – it is open ended. Considers topics relevant to more developed countries. All issues brought in which may affect human beings' potential.
3. Freedom to choose given priority – I.e. HD concerns **widening human choices**.
4. Incomes are means not end.
5. Yet accepts that humans are an important resource too as well as being the objective of development.

Advantages compared with BN and capability approaches

- Goes beyond 'basic'. Much better to have an approach that encompasses all nations and income levels.
- Goes beyond physical condition to institutional and political elements.
- Tries to add up and assess country progress. Here better than capability approach.
- To some extent a political agenda, work-in-progress, a rallying cry for all those seeking human and humane alternatives, evaluating our current condition.

Thank You