Structural Equation Modeling-
A Simple-complex Multivariate technique

Prabhat Mittal
profmittal@yahoo.co.in

Objectives

- ✓ Basic concepts in SEM
- ✓ Statistics associated with SEM
- ✓ Foundations of SEM
- ✓ Process of conducting SEM
- ✓ How to specify the measurement model
- ✓ Assess Measurement model Reliability and Validity
- ✓ Specify the structural model
- ✓ Assess structural model validity
- ✓ Relationship of SEM to other Multivariate techniques

Basic Concepts in SEM

Basic Concept

- Marketing research must answer a set of interrelated questions-
  - What variables determine service quality?
    - How does service quality influence service attitude and service satisfaction?
      - How does satisfaction with the service result in patronage intention?
      - How does attitude toward the service combine with other variables to affect intention to patronize the service?

Instructor: Dr. Prabhat Mittal M. Sc., M.Phil, Ph.D. (FMS, DU)
Post-doctoral, University of Minnesota, USA
E-mail Id: profmittal@yahoo.co.in

Basic Concept

- SEM examines the structure of these interrelationships, which are expressed in a series of structural equations similar to estimating a series of multiple regression equations between one or more Independent Variables (IV) and one or more Dependent Variables (DV).
- These equations model all the relationships among dependent as well as independent constructs (unobservable or latent factors) represented by the observed variables similar to the concept of variables representing a factor in factor analysis.

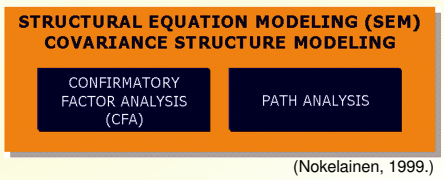
Basic Concept

- SEM or *Covariance structure analysis*, *latent variable analysis*, and *causal modeling*.
- SEM is mainly used as a confirmatory, rather than exploratory technique.
  - It determine whether a certain model is valid, rather than using SEM to "find" a suitable model.
- SEM can not establish causality, although it can assist in that process.

Components of SEM

- Two main components of SEM are presented in Figure 1.
  - CFA operates with observed and latent variables, path analysis operates only with observed variables.

Figure 1. Components of Structural Equation Modeling



(Nokelainen, 1999.)

Basic Concepts of Factor Analysis

- The fundamental idea underlying the factor analysis is that some but not all variables can be directly observed.
- Those unobserved variables are referred to as either *latent* variables or factors.
- Information about latent variables can be gained by observing their influence on observed variables.
- Factor analysis examines covariation among a set of observed variables trying to *generate a smaller number of latent variables*.

Instructor: **Dr. Prabhat Mittal** M. Sc., M.Phil, Ph.D. (FMS, DU)
Post-doctoral, University of Minnesota, USA
E-mail Id: profmittal@yahoo.co.in

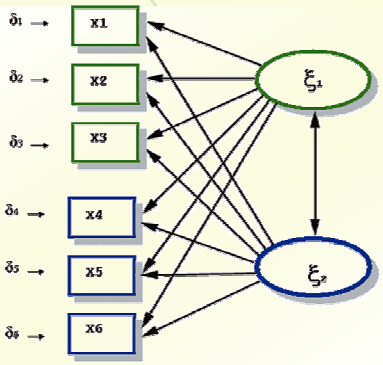
Basic Concepts of Factor Analysis

- ***Exploratory Factor Analysis***
  - In exploratory factor analysis (EFA), observed variables are represented by squares and circles represent latent variables.
  - Causal effect of the latent variable on the observed variable is presented with straight line with arrowhead.

Basic Concepts of Factor Analysis

- ***Exploratory Factor Analysis***
  - The latent factors (ellipses) labeled with ξ's (Xi) are called common factors and the δ's (delta) (usually in circles) are called errors in variables or *residual variables*.
  - Errors in variables have unique effects to one and only one observed variable - unlike the common factors that share their effects in common with more than one of the observed variables.

Basic Concepts of Factor Analysis



Basic Concepts of Factor Analysis

- ***Exploratory Factor Analysis***
  - The EFA model reflects the fact that researcher does not specify the structure of the relationships among the variables in the model.
  - When carrying out EFA, researcher must assume that
    - all common factors are correlated,
    - all observed variables are directly affected by all common factors,
    - errors in variables are uncorrelated with one another,
    - all observed variables are affected by a unique factor and
    - all ξ's are uncorrelated with all δ's. (Long, 1983.)

Instructor: **Dr. Prabhat Mittal** M. Sc., M.Phil, Ph.D. (FMS, DU)
Post-doctoral, University of Minnesota, USA
E-mail Id: profmittal@yahoo.co.in

Basic Concepts of Factor Analysis

- ***Confirmatory Factor Analysis***
  - One of the biggest problems in EFA is its inability to incorporate substantively meaningful constraints.
  - That is due to fact that algebraic mathematical solution to solve estimates is not trivial, instead one has to seek for other solutions.
  - That problem was partly solved by the development of the confirmatory factor model, which was based on an iterative algorithm (Jöreskog, 1969).

Basic Concepts of Factor Analysis

- ***Confirmatory Factor Analysis***
  - In confirmatory factor analysis (CFA), which is a special case of SEM, the correlations between the factors are an explicit part of the analysis because they are collected in a matrix of factor correlations.
  - With CFA, researcher is able to decide *a priori* whether the factors would correlate or not. (Tacq, 1997.)

Basic Concepts of Factor Analysis

- ***Confirmatory Factor Analysis***
  - Moreover, researcher is able to impose substantively motivated constraints,
    - which common factor pairs that are correlated,
    - which observed variables are affected by which common factors,
    - which observed variables are affected by a unique factor and
    - which pairs of unique factors are correlated. (Long, 1983.)

Path Analysis

- Examines how *n* independent (x, IV, Xi, ξ) variables are statistically related to a dependent (y, DV, Eta, η) variable.
- Applies the techniques of regression analysis, aiming at more detailed resolution of the phenomena under investigation.
- Allows
  - *Causal* interpretation of statistical dependencies
  - Examination of how data fits to a theoretical model

Instructor: **Dr. Prabhat Mittal** M. Sc., M.Phil, Ph.D. (FMS, DU)
Post-doctoral, University of Minnesota, USA
E-mail Id: profmittal@yahoo.co.in

Path Analysis

- Once the data is available, conduction of path analysis is straightforward:
  1. Draw a path diagram according to the theory.
  2. Conduct one or more regression analyses.
  3. Compare the regression estimates (B) to the theoretical assumptions or (Beta) other studies.
  4. If needed, modify the model by removing or adding connecting paths between the variables and redo stages 2 and 3.

Path Analysis

- Data assumptions:
  - DV:
    - Continuous, normally distributed (univariate normality assumption)
  - IV:
    - Continuous (no dichotomy or categorical variables)
  - N:
    - About 30 observations for each IV

Path Analysis

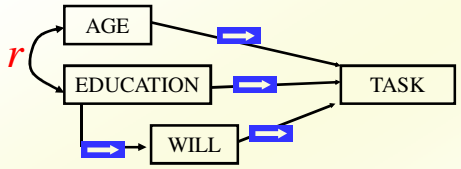
- Theoretical assumptions
  - Causality:
    - X_1 and Y_1 correlate.
    - X_1 precedes Y_1 chronologically.
    - X_1 and Y_1 are still related after controlling other dependencies.
- Statistical assumptions
  - Model needs to be recursive.
  - It is OK to use ordinal data.
  - All variables are measured (and analyzed) without measurement error ($\varepsilon = 0$).

Path Analysis

- As stated earlier, path analysis assumes that the model is recursive.
  - Nature of causal dependency is unidirectional, like a 'one way road' (arc with one head →).
  - If there is no a priori information available about the direction of causal dependency, it is assumed to be correlational (arc with two heads ↔).

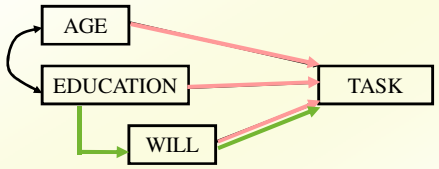
Instructor: **Dr. Prabhat Mittal** M. Sc., M.Phil, Ph.D. (FMS, DU)
Post-doctoral, University of Minnesota, USA
E-mail Id: profmittal@yahoo.co.in

Path Analysis



Path Analysis

- Direct and indirect effect



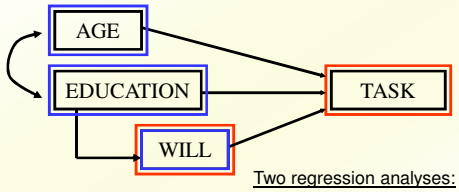
Path Analysis

- There are two types of observed variables:
  - Endogenous (y, DV, Eta η) → DV
  - Exogenous (x, IV, Xi ξ). → IV
- For each endogenous (DV) variable, a regression analysis is performed.

Path Analysis

x IV Xi ξ EXOGENIOUS

y DV Eta η ENDOGENIOUS



Two regression analyses:

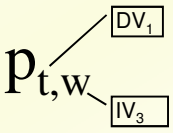
1) AGE + EDUCATION + WILL -> TASK

2) EDUCATION –> WILL

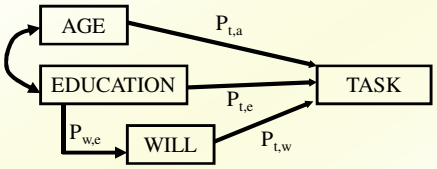
Instructor: **Dr. Prabhat Mittal** M. Sc., M.Phil, Ph.D. (FMS, DU)
Post-doctoral, University of Minnesota, USA
E-mail Id: profmittal@yahoo.co.in

Path Analysis

- Path coefficients are a product of one or more regression analyses.
  - They are indicators of statistical dependency between variables.



Path Analysis



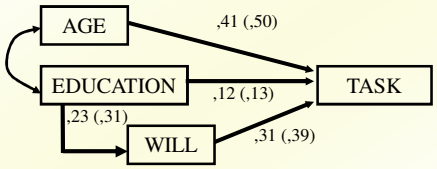
Path Analysis

- Path coefficients are standardized ('Beta') or unstandardized ('B') regression coefficients.
  - Strength of inter-variable dependencies are comparable to other studies when standardized values (z, where $M = 0$ and $SD = 1$) are used.
  - Unstandardized values allow the original measurement scale examination of inter-variable dependencies.

$$SD = \sqrt{\frac{\sum (x - \bar{x})^2}{N - 1}} \qquad z = \frac{(x - \bar{x})}{SD}$$

Path Analysis

- Beta (B)



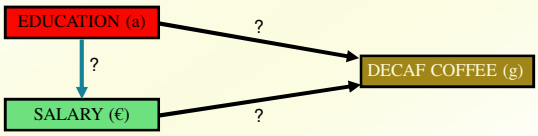
Instructor: **Dr. Prabhat Mittal** M. Sc., M.Phil, Ph.D. (FMS, DU)
Post-doctoral, University of Minnesota, USA
E-mail Id: profmittal@yahoo.co.in

Path Analysis

- Path coefficient ($p_{DV,IV}$) indicates the direct effect of IV to DV.
- If the model contains only one IV and DV variable, the path coefficient equals to correlation coefficient.
  - In those models that have more than two variables (one IV and one DV), the path coefficients equal to partial correlation coefficients.
    - The other path coefficients are controlled while each individual path coefficient is calculated.

Path Analysis

- No need to use LISREL or AMOS
  - Two separate regression analyses in SPSS (Analyze – Regression – Linear)



1. Data ($N = 10$)

	EDUCATION	SALARY	DECAF_COFFEE
1	20	4000	500
2	18	3600	600
3	12	700	0
4	8	200	0
5	21	5000	1000
6	9	3000	0
7	18	1400	360
8	17	1200	200
9	8	1000	0
10	9	800	0

2. First SPSS regression analysis (SALARY + EDUCATION -> DECAF_COFFEE)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,939[a]	,877	,843	136,969

a. Predictors: (Constant), SALARY, EDUCATION

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-447,534	137,286		-3,260	,014
	EDUCATION	33,224	11,752	,506	2,827	,026
	SALARY	,108	,037	,518	2,894	,023

a. Dependent Variable: DECAF_COFFEE

3. Second SPSS regression analysis (EDUCATION -> SALARY)

Model Summary

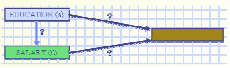
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,673[a]	,453	,386	1295,232

a. Predictors: (Constant), EDUCATION

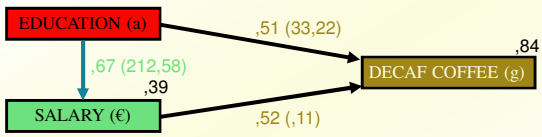
Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-981,169	1256,814		-,781	,457
	EDUCATION	212,581	82,514	,673	2,576	,033

a. Dependent Variable: SALARY



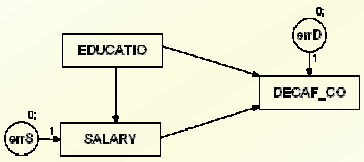
Path Analysis



Instructor: **Dr. Prabhat Mittal** M. Sc., M.Phil, Ph.D. (FMS, DU)
Post-doctoral, University of Minnesota, USA
E-mail Id: profmittal@yahoo.co.in

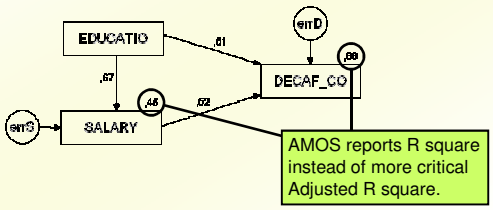
Path Analysis

- Here is the same model in AMOS:



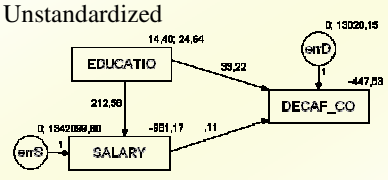
Path Analysis

- And the results are naturally the same:
  - Standardized



Path Analysis

- And the results are naturally the same:
  - Unstandardized



When SEM is Useful?

- When SEM is useful:
  - When you need to deal with latent (unobserved) constructs.
  - When you have a strong theoretical background to your data (a priori hypothesis).
  - When you are examining complex relationships.
  - When you have access to a large sample.

Instructor: **Dr. Prabhat Mittal** M. Sc., M.Phil, Ph.D. (FMS, DU)
Post-doctoral, University of Minnesota, USA
E-mail Id: profmittal@yahoo.co.in

When SEM is Useful?

- **Social Psychology**
  - Structural equation modeling the use of a risk assessment instrument in child protective services
  - A Structural Equation Model of Social Influences and Exposure to Media Smoking on Adolescent Smoking
- **Business/ Commerce**
  - Application of structural equation modeling to evaluate the Intention of shippers to use Internet services in liner shipping.
  - Human Capital and SME Internationalization: A Structural Equation Modeling Study
- **Health/Medicine**
  - Application of Structural Equation Modeling to Health Outcomes Research
  - Structural equation modeling of sleep apnea, inflammation, and metabolic dysfunction in children

When SEM is Useful?

- **Neurosciences**
  - Connectivity exploration with structural equation modeling: an fMRI study of bimanual motor coordination
  - Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data
  - A structural equation modeling analysis of attentional control: an event-related fMRI study
- **Cognition**
  - Validation of Cognitive Structures: A Structural Equation Modeling Approach.
  - Static and Dynamic Longitudinal Structural Analyses of Cognitive Changes in Old Age

Foundations of SEM

Defining variables in research

- Variables-(construct) it denotes a symbol of an event, act, characteristic, trait or attribute that can be measured and to which we assign categorical values.
  - For purpose of data entry and analysis, numerical values like 0,1 is assigned to male-female, employed-unemployed etc..
- Independent (IV) vs. dependent variable (DV)- Predicator variable or IV is the variable that causes an effect on DV. While criteria variable or DV is measured, predicted expected to be affected by the IV.
- Moderating variable (MV)- it is a second type of IV that is included because it is believed to have a significant contributory or contingent effect of the original stated IV-DV relationship.
  - For e.g. Salary compensation (IV) leading to sales productivity (DV), especially among younger workers (MV)

Instructor: **Dr. Prabhat Mittal** M. Sc., M.Phil, Ph.D. (FMS, DU)
Post-doctoral, University of Minnesota, USA
E-mail Id: profmittal@yahoo.co.in

Extraneous Variables

Continued

- Extraneous variables (EV) – These are IV(s)s not related to the purpose of the study, but have a insignificant effect on DV and hence excluded from the study. Whatever effect is noticed on DV is technically described as an 'experimental error'.
- Sometimes the extraneous variables-control variable is used to minimize the effects of EV. It refers to restrain experimental conditions.
  - For e.g. with new customers (EV-control variable), Salary compensation (IV) leading to increased sales productivity (DV), especially among younger workers (MV).
- When the DV is not free from the influence of EV(s), the relationship between the DV and IV is said to be confounded by an EV(s).

Intervening Variables

Continued

- Intervening variables (IVV)-factor which theoretically affects the observed phenomenon but cannot be seen, measured, or manipulated; its effect must be inferred from the effects of the IV and MV on the observed phenomenon.
  - For e.g. with new customers (EV-control variable) introduction of training program (IVV), Salary compensation (IV) leading to increased sales productivity (DV), especially among younger workers (MV).

Concepts vs Constructs

- Concepts-generally accepted collection of meanings of characteristics associated with the current events, objects, conditions, situations, and behaviors. For e.g. we see a man passing and identify his attitude like running, walking, skipping, crawling or hopping represent concepts.
- Constructs (*factor*)- is an image or abstract idea specifically invented for a given research/theory-building purpose. Combining concrete concepts, *when the idea can not be directly observed builds constructs*

Problems with Mobile operator combination of concepts-
- Errors in mobile bill
- Extra charges
- Network problem

Most concrete and can be easily measured

Customer Satisfaction with Mobile Operator (where components are unknown)

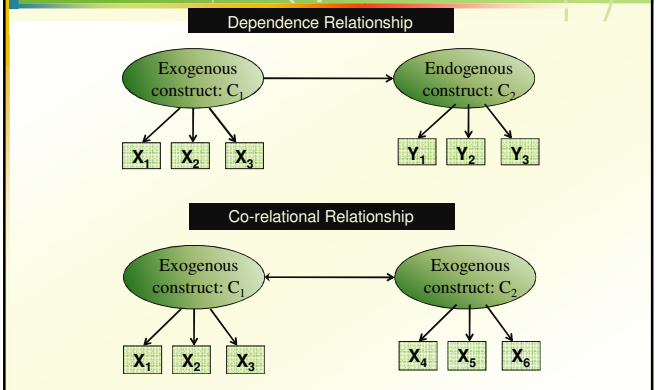
These type of constructs are hypothetical constructs, which can be inferred only from the data. They are presumed to exist but must await further testing.

Exogenous vs. Endogenous Constructs

- Exogenous construct- acts as an independent variable represented by a multiple observed variables outside the model. It cannot be explained by any other construct or variable in the model.
  - Graphically, it does not have any paths coming into it from any construct or variable in the model; it will only have paths (single headed arrows) going out of it.
- Endogenous construct-latent, multi-item equivalent of a dependent variable. It is dependent on other constructs.
  - Graphically, it has one or more paths coming into it from one or more exogenous or endogenous constructs. Indicators of endogenous are referred to as 'Y' variable.

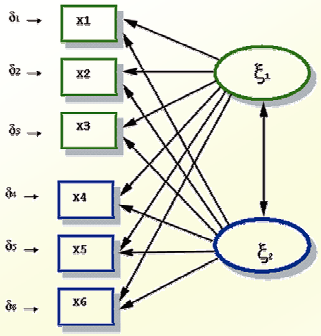
Instructor: **Dr. Prabhat Mittal** M. Sc., M.Phil, Ph.D. (FMS, DU)
Post-doctoral, University of Minnesota, USA
E-mail Id: profmittal@yahoo.co.in

Dependence vs Co-relational relationships



Prabhat Mittal
profmittal@yahoo.co.in

Path Diagram



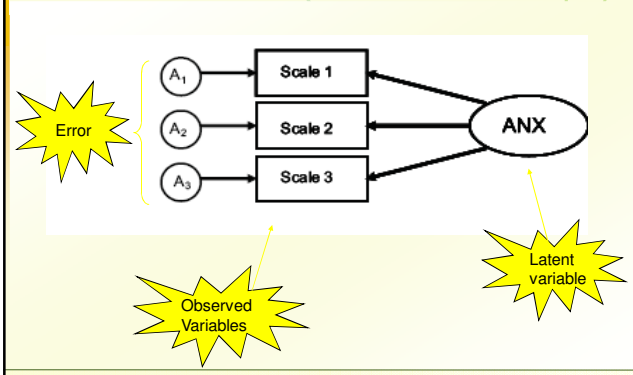
(Nokelainen, 1999.)

Note:

- Constructs represented by oval/circle
- Measured variables by squares
- Straight arrows are drawn from the constructs to the measure variables.
- Dependence relations by single headed straight arrows
- Co-relational relationships by two-side headed straight/curved arrows.

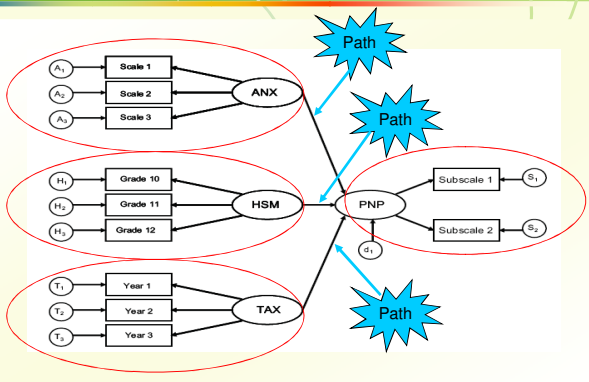
Prabhat Mittal
profmittal@yahoo.co.in

CFA diagram



Prabhat Mittal
profmittal@yahoo.co.in

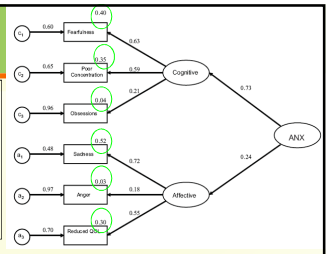
Complete SEM diagram



Prabhat Mittal
profmittal@yahoo.co.in

Instructor: **Dr. Prabhat Mittal** M. Sc., M.Phil, Ph.D. (FMS, DU)
Post-doctoral, University of Minnesota, USA
E-mail Id: profmittal@yahoo.co.in

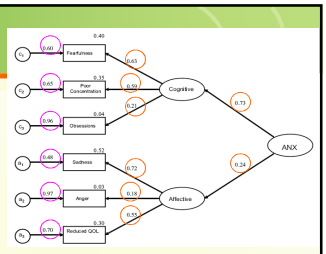
Squared values of the path coefficient (SMR)



SMR = Squared value of path coefficients

- Interpreted like an R^2 multiple regression
  - in terms of how much of the variance in one variable is explained by, or is in common with, the other variable.

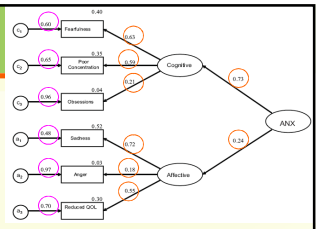
Error variances



Error variances = # over the arrows between the error terms and the observed variables.

- NB: Σ (SMR + error variance per variable) = 1
  - all the variance of a variable is divided between that shared with the latent variable and error.

Path coefficients



Path coefficient is equivalent to the factor loadings in FA.

- Therefore, this is a regression value.
  - Standardized coefficients range: -1 to 1
  - "> value" = stronger association

Proposition and Hypotheses

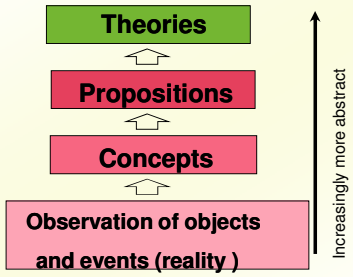
- Proposition-statement about observable phenomena (concept) that may be judged as true or false. When a proposition is formulated for empirical testing, it is called as Hypothesis. The description of the hypothesis leads to a case.
- Type of Hypothesis
  - Descriptive Hypothesis-state the existence, size, form or distribution of some variable. For e.g.
    Average monthly cell phone bills of this city is more than Rs.700
    Average no. of TV sets in Indian homes is 3
  - Relational Hypothesis-describes a relationship between two variables with respect to some case. For e.g.
    Co-relational Hypothesis:
    Average monthly cell phone bills of Delhi is more than that of Mumbai
    Explanatory (causal) Hypothesis:
    Increase in family income (IV) leads to an increase in the percentage of income saved (DV)

Instructor: **Dr. Prabhat Mittal** M. Sc., M.Phil, Ph.D. (FMS, DU)
Post-doctoral, University of Minnesota, USA
E-mail Id: profmittal@yahoo.co.in

Theory

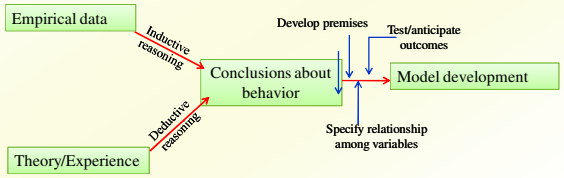
- A theory is a set of systematically interrelated concepts, definitions, and propositions that are advanced to explain and predict phenomena (facts). In this sense, many theories are developed and continually used to explain or predict what goes on around us.
- To the degree that our theories are sound and fit the situation leads to better explanations and predictions of the situation.
- In general, theories tend to be complex, be abstract, and involve multiple variables. While hypotheses, tend to be more simple, limited-variable statements involving concrete instances.

Theory Building A Process Of Increasing Abstraction



Model Development

- Model used in business research and other fields to represent phenomena through the use of analogy.
- It is a representation of a system that is constructed to study some aspect of that system or the system as a whole.
- Models differ from theories in that a theory's role is explanation whereas a model's role is representation.



SEM Model

Instructor: Dr. Prabhat Mittal M. Sc., M.Phil, Ph.D. (FMS, DU)
Post-doctoral, University of Minnesota, USA
E-mail Id: profmittal@yahoo.co.in