A Brief History of Scientific Thoughts
Lecture 3

Palash Sarkar

Applied Statistics Unit
Indian Statistical Institute, Kolkata
India
palash@isical.ac.in



Observation and Inference

- Observation: people pass away.
- Inference: all men are mortal.

- Observation: movements of the planets by Tycho Brahe.
- Inference: planets move in elliptical orbits with the Sun at one of the foci (Johannes Kepler).

- Observation: one million triangles whose sum-of-angles are the same.
- Inference: the sum of angles of any triangle is a constant.



Observation and Inference

- Observation: people pass away.
- Inference: all men are mortal.

- Observation: movements of the planets by Tycho Brahe.
- Inference: planets move in elliptical orbits with the Sun at one of the foci (Johannes Kepler).

- Observation: one million triangles whose sum-of-angles are the same.
- Inference: the sum of angles of any triangle is a constant.



Observation and Inference

- Observation: people pass away.
- Inference: all men are mortal.

- Observation: movements of the planets by Tycho Brahe.
- Inference: planets move in elliptical orbits with the Sun at one of the foci (Johannes Kepler).

- Observation: one million triangles whose sum-of-angles are the same.
- Inference: the sum of angles of any triangle is a constant.



Observation and Inference

- Observation: people pass away.
- Inference: all men are mortal.

- Observation: movements of the planets by Tycho Brahe.
- Inference: planets move in elliptical orbits with the Sun at one of the foci (Johannes Kepler).

- Observation: one million triangles whose sum-of-angles are the same.
- Inference: the sum of angles of any triangle is a constant.



Observation and Inference

- Observation: people pass away.
- Inference: all men are mortal.

- Observation: movements of the planets by Tycho Brahe.
- Inference: planets move in elliptical orbits with the Sun at one of the foci (Johannes Kepler).

- Observation: one million triangles whose sum-of-angles are the same.
- Inference: the sum of angles of any triangle is a constant.



Observation and Inference

- Observation: people pass away.
- Inference: all men are mortal.

- Observation: movements of the planets by Tycho Brahe.
- Inference: planets move in elliptical orbits with the Sun at one of the foci (Johannes Kepler).

- Observation: one million triangles whose sum-of-angles are the same.
- Inference: the sum of angles of any triangle is a constant.



Observation and Inference

- Observation: people pass away.
- Inference: all men are mortal.

- Observation: movements of the planets by Tycho Brahe.
- Inference: planets move in elliptical orbits with the Sun at one of the foci (Johannes Kepler).

- Observation: one million triangles whose sum-of-angles are the same.
- Inference: the sum of angles of any triangle is a constant.



Inductive Inference

The type of inference in each of the above cases is different from the type of inference in the following argument.

Given:
All men are mortal;
Socrates is a man;
Inference:
therefore Socrates is mortal.

- Former kind: inductive inference/logic/reasoning/procedure.
- Later kind: deductive inference/logic/reasoning/procedure.
- Scientific inferences are usually (if not all) inductive.
- The method of induction as used in mathematics is a deductive procedure.
- Temper of mind.
  - Axiomatic-deductive: mathematics/logic.
  - Inductive inference: science.



Inductive Inference

The type of inference in each of the above cases is different from the type of inference in the following argument.

Given:
All men are mortal;
Socrates is a man;
Inference:
therefore Socrates is mortal.

- Former kind: inductive inference/logic/reasoning/procedure.
- Later kind: deductive inference/logic/reasoning/procedure.
- Scientific inferences are usually (if not all) inductive.
- The method of induction as used in mathematics is a deductive procedure.
- Temper of mind.
  - Axiomatic-deductive: mathematics/logic.
  - Inductive inference: science.



Inductive Inference

The type of inference in each of the above cases is different from the type of inference in the following argument.

Given:
All men are mortal;
Socrates is a man;
Inference:
therefore Socrates is mortal.

- Former kind: inductive inference/logic/reasoning/procedure.
- Later kind: deductive inference/logic/reasoning/procedure.
- Scientific inferences are usually (if not all) inductive.
- The method of induction as used in mathematics is a deductive procedure.
- Temper of mind.
  - Axiomatic-deductive: mathematics/logic.
  - Inductive inference: science.



Inductive Inference

The type of inference in each of the above cases is different from the type of inference in the following argument.

Given:
All men are mortal;
Socrates is a man;
Inference:
therefore Socrates is mortal.

- Former kind: inductive inference/logic/reasoning/procedure.
- Later kind: deductive inference/logic/reasoning/procedure.
- Scientific inferences are usually (if not all) inductive.
- The method of induction as used in mathematics is a deductive procedure.
- Temper of mind.
  - Axiomatic-deductive: mathematics/logic.
  - Inductive inference: science.



Induction from Simple Enumeration

A census officer who had to record the names of males in a Welsh village.

- The first was called William Williams.
- The second was called William Williams.
- The third, fourth, ... were called William Williams.
- The officer decides that all are called William Williams.
- Problem (hasty generalisation): He was wrong. One was called John Johnes.

A parable from “History of Western Philosophy” by Bertrand Russell.



Induction from Simple Enumeration

A census officer who had to record the names of males in a Welsh village.

- The first was called William Williams.
- The second was called William Williams.
- The third, fourth, ... were called William Williams.
- The officer decides that all are called William Williams.
- Problem (hasty generalisation): He was wrong. One was called John Johnes.

A parable from “History of Western Philosophy” by Bertrand Russell.



Induction from Simple Enumeration

A census officer who had to record the names of males in a Welsh village.

- The first was called William Williams.
- The second was called William Williams.
- The third, fourth, ... were called William Williams.
- The officer decides that all are called William Williams.
- Problem (hasty generalisation): He was wrong. One was called John Johnes.

A parable from “History of Western Philosophy” by Bertrand Russell.



Induction from Simple Enumeration

A census officer who had to record the names of males in a Welsh village.

- The first was called William Williams.
- The second was called William Williams.
- The third, fourth, ... were called William Williams.
- The officer decides that all are called William Williams.
- Problem (hasty generalisation): He was wrong. One was called John Johnes.

A parable from “History of Western Philosophy” by Bertrand Russell.



Induction from Simple Enumeration

A census officer who had to record the names of males in a Welsh village.

- The first was called William Williams.
- The second was called William Williams.
- The third, fourth, ... were called William Williams.
- The officer decides that all are called William Williams.
- Problem (hasty generalisation): He was wrong. One was called John Johnes.

A parable from “History of Western Philosophy” by Bertrand Russell.



Induction from Simple Enumeration

A census officer who had to record the names of males in a Welsh village.

- The first was called William Williams.
- The second was called William Williams.
- The third, fourth, ... were called William Williams.
- The officer decides that all are called William Williams.
- Problem (hasty generalisation): He was wrong. One was called John Johnes.

A parable from “History of Western Philosophy” by Bertrand Russell.



Induction from Simple Enumeration

A census officer who had to record the names of males in a Welsh village.

- The first was called William Williams.
- The second was called William Williams.
- The third, fourth, ... were called William Williams.
- The officer decides that all are called William Williams.
- Problem (hasty generalisation): He was wrong. One was called John Johnes.

A parable from “History of Western Philosophy” by Bertrand Russell.



Empiricism in Science

Sir Francis Bacon (1561–1626)

- Founder of the modern inductive method and pioneer at systematization of scientific procedure.
- Commonly regarded as the originator of the saying ‘Knowledge is power’.

Contributions:

- Rejected simple enumeration and suggested a method of induction.
- Undervalued mathematics as insufficiently experimental.
- Idols: bad mental habits which cause errors.



Empiricism in Science

Sir Francis Bacon (1561–1626)

- Founder of the modern inductive method and pioneer at systematization of scientific procedure.
- Commonly regarded as the originator of the saying ‘Knowledge is power’.

Contributions:

- Rejected simple enumeration and suggested a method of induction.
- Undervalued mathematics as insufficiently experimental.
- Idols: bad mental habits which cause errors.



Empiricism in Science

Sir Francis Bacon (1561–1626)

- Founder of the modern inductive method and pioneer at systematization of scientific procedure.
- Commonly regarded as the originator of the saying ‘Knowledge is power’.

Contributions:

- Rejected simple enumeration and suggested a method of induction.
- Undervalued mathematics as insufficiently experimental.
- Idols: bad mental habits which cause errors.



Bacon's Induction

Suppose the intention is to discover the nature of heat.

Observation:

- Draw up a list of hot bodies.
- Draw up a list of cold bodies.
- Draw up a list of bodies of varying degrees of heat.

Inference:

- The lists would show characteristics always present in hot bodies, absent in cold bodies and present in varying degrees in bodies of different degrees of heat.
- By this method one arrives at a general law of the lowest degree of generality.
- From such laws one arrives at laws of the second degree of generality, and so on.



Bacon's Induction

Suppose the intention is to discover the nature of heat.
Observation:

- Draw up a list of hot bodies.
- Draw up a list of cold bodies.
- Draw up a list of bodies of varying degrees of heat.

Inference:

- The lists would show characteristics always present in hot bodies, absent in cold bodies and present in varying degrees in bodies of different degrees of heat.
- By this method one arrives at a general law of the lowest degree of generality.
- From such laws one arrives at laws of the second degree of generality, and so on.



Bacon's Idols

- Idols of the tribe: those inherent in human nature;
  e.g.: expecting more order in nature than actual.
- Idols of the cave: personal prejudices which are characteristic of the investigator.
- Idols of the market-place: confusion in the use of words and the difference in the scientific meaning versus common usage.
- Idols of the theatre: mindset and dogma developing out of received systems of thoughts and not asking questions.



Bacon's Idols

- Idols of the tribe: those inherent in human nature;
  e.g.: expecting more order in nature than actual.
- Idols of the cave: personal prejudices which are characteristic of the investigator.
- Idols of the market-place: confusion in the use of words and the difference in the scientific meaning versus common usage.
- Idols of the theatre: mindset and dogma developing out of received systems of thoughts and not asking questions.



Bacon's Idols

- Idols of the tribe: those inherent in human nature;
  e.g.: expecting more order in nature than actual.
- Idols of the cave: personal prejudices which are characteristic of the investigator.
- Idols of the market-place: confusion in the use of words and the difference in the scientific meaning versus common usage.
- Idols of the theatre: mindset and dogma developing out of received systems of thoughts and not asking questions.



Bacon's Idols

- Idols of the tribe: those inherent in human nature;
  e.g.: expecting more order in nature than actual.
- Idols of the cave: personal prejudices which are characteristic of the investigator.
- Idols of the market-place: confusion in the use of words and the difference in the scientific meaning versus common usage.
- Idols of the theatre: mindset and dogma developing out of received systems of thoughts and not asking questions.



Criticisms of Bacon

- Insufficient emphasis on hypothesis:
  - A mere orderly arrangement of the data does not make the hypothesis obvious.
  - Formulation of hypothesis is one of the main scientific tasks.
  - Some hypothesis is necessary to even start collection of data.
- Role of deduction in science:
  - Testing: a long journey from a hypothesis to a consequence which can be tested by observation.
  - This journey is deductive.
- Issue of induction by simple enumeration.
  - May be rejected when scientific enumeration is concerned.
  - We may assume more general laws.
  - But, the validity of these laws are themselves based on simple enumeration.



Criticisms of Bacon

- Insufficient emphasis on hypothesis:
  - A mere orderly arrangement of the data does not make the hypothesis obvious.
  - Formulation of hypothesis is one of the main scientific tasks.
  - Some hypothesis is necessary to even start collection of data.
- Role of deduction in science:
  - Testing: a long journey from a hypothesis to a consequence which can be tested by observation.
  - This journey is deductive.
- Issue of induction by simple enumeration.
  - May be rejected when scientific enumeration is concerned.
  - We may assume more general laws.
  - But, the validity of these laws are themselves based on simple enumeration.



Criticisms of Bacon

- Insufficient emphasis on hypothesis:
  - A mere orderly arrangement of the data does not make the hypothesis obvious.
  - Formulation of hypothesis is one of the main scientific tasks.
  - Some hypothesis is necessary to even start collection of data.
- Role of deduction in science:
  - Testing: a long journey from a hypothesis to a consequence which can be tested by observation.
  - This journey is deductive.
- Issue of induction by simple enumeration.
  - May be rejected when scientific enumeration is concerned.
  - We may assume more general laws.
  - But, the validity of these laws are themselves based on simple enumeration.



Mill's Method of Induction

John Stuart Mill (1806–1873):

- Formulated five laws of causality.
- The first three follow from Bacon's work on induction.

`http://en.wikipedia.org/wiki/Mill's_Methods`



Direct Method of Agreement

"If two or more instances of the phenomenon under investigation have only one circumstance in common, the circumstance in which alone all the instances agree, is the cause (or effect) of the given phenomenon."

- For a property to be a necessary condition it must always be present if the effect is present.
- Symbolically, the method of agreement can be represented as:
  A B C D occur together with w x y z
  A E F G occur together with w t u v;
  therefore A is the cause, the effect, or part of the cause of w.



Direct Method of Agreement

"If two or more instances of the phenomenon under investigation have only one circumstance in common, the circumstance in which alone all the instances agree, is the cause (or effect) of the given phenomenon."

- For a property to be a necessary condition it must always be present if the effect is present.
- Symbolically, the method of agreement can be represented as:
  A B C D occur together with w x y z
  A E F G occur together with w t u v;
  therefore A is the cause, the effect, or part of the cause of w.



Direct Method of Agreement

"If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance in common save one, that one occurring only in the former; the circumstance in which alone the two instances differ, is the effect, or the cause, or an indispensable part of the cause, of the phenomenon."

- Symbolically, the method of difference can be represented as:
  A B C D occur together with w x y z
  B C D occur together with x y z;
  therefore A is the cause, or the effect, or a part of the cause of w.



Direct Method of Agreement

"If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance in common save one, that one occurring only in the former; the circumstance in which alone the two instances differ, is the effect, or the cause, or an indispensable part of the cause, of the phenomenon."

- Symbolically, the method of difference can be represented as:
  A B C D occur together with w x y z
  B C D occur together with x y z;
  therefore A is the cause, or the effect, or a part of the cause of w.



Joint Method of Agreement and Difference

"If two or more instances in which the phenomenon occurs have only one circumstance in common, while two or more instances in which it does not occur have nothing in common save the absence of that circumstance: the circumstance in which alone the two sets of instances differ, is the effect, or cause, or a necessary part of the cause, of the phenomenon."

- This principle simply represents the application of the methods of agreement and difference.
- Symbolically, the joint method of agreement and difference can be represented as:
  A B C occur together with x y z
  A D E occur together with x v w also B C occur with y z;
  therefore A is the cause, or the effect, or a part of the cause of X.



Joint Method of Agreement and Difference

"If two or more instances in which the phenomenon occurs have only one circumstance in common, while two or more instances in which it does not occur have nothing in common save the absence of that circumstance: the circumstance in which alone the two sets of instances differ, is the effect, or cause, or a necessary part of the cause, of the phenomenon."

- This principle simply represents the application of the methods of agreement and difference.
- Symbolically, the joint method of agreement and difference can be represented as:
  A B C occur together with x y z
  A D E occur together with x v w also B C occur with y z;
  therefore A is the cause, or the effect, or a part of the cause of x.



Method of Residues

"Deduct from any phenomenon such part as is known by previous inductions to be the effect of certain antecedents, and the residue of the phenomenon is the effect the remaining antecedents."

- Symbolically, the method of residues can be represented as:
  A B C occur together with x y z
  B is known to be the cause of y
  C is known to be the cause of z;
  therefore A is the cause or effect of x.



Method of Residues

"Deduct from any phenomenon such part as is known by previous inductions to be the effect of certain antecedents, and the residue of the phenomenon is the effect the remaining antecedents."

- Symbolically, the method of residues can be represented as:
  A B C occur together with x y z
  B is known to be the cause of y
  C is known to be the cause of z;
  therefore A is the cause or effect of x.



Method of Concomitant Variations

"Whatever phenomenon varies in any manner whenever another phenomenon varies in some particular manner, is either a cause or an effect of that phenomenon, or is connected with it through some fact of causation."

- Samples of toxic water containing salt and lead; as proportion of lead increases/decreases, the toxicity increases/decreases.
- Symbolically, the method of concomitant variation can be represented as:
  A B C occur together with x y z
  (A+ or A-) B C result in (x+ or x-) y z;
  therefore A and x are causally connected.



Method of Concomitant Variations

"Whatever phenomenon varies in any manner whenever another phenomenon varies in some particular manner, is either a cause or an effect of that phenomenon, or is connected with it through some fact of causation."

- Samples of toxic water containing salt and lead; as proportion of lead increases/decreases, the toxicity increases/decreases.
- Symbolically, the method of concomitant variation can be represented as:
  A B C occur together with x y z
  (A+ or A-) B C result in (x+ or x-) y z;
  therefore A and x are causally connected.



Auguste Comte (1798–1857)

- Founder of a philosophical movement called positivism.
  - Only authentic knowledge is that which is based on sense experience and positive verification.
  - Different forms of the positive method: astronomy–observation, physics–experimentation, biology–comparison.
  - Later variants: post-positivism, logical positivism, etcetera.
- Law of three stages.
  - Humanity passes through three successive stages: the theological, the metaphysical, and the positive (final).



Auguste Comte (1798–1857)

- Founder of a philosophical movement called positivism.
  - Only authentic knowledge is that which is based on sense experience and positive verification.
  - Different forms of the positive method: astronomy–observation, physics–experimentation, biology–comparison.
  - Later variants: post-positivism, logical positivism, etcetera.
- Law of three stages.
  - Humanity passes through three successive stages: the theological, the metaphysical, and the positive (final).



Auguste Comte (1798–1857) (contd.)

Considered to be the founder of the philosophy of science.

- First to observe the circularity in theory and observation.
- Classification of the six sciences.
- Introduction of the scientific method to sociology.
  - Argued that development of natural sciences was a necessary prelude to the development of social science.



Theory-Observation Circularity

- Observation involves both perception and cognition.
- Depends on our conception of the world; that conception may influence what is perceived or considered worthy of perception.
- Most scientific observation must be done within a theoretical context in order to be useful.



Theory-Observation Circularity (contd.)

Empirical observation and acceptability of a hypothesis.

- The observation is framed in terms of a theory that also contains the hypothesis it is meant to verify or falsify.
- So, the observation can only arbitrate between the hypotheses within the context of the underlying theory.

Inductive bias.

- A modern notion used for learning algorithms.
- Assumptions that are used to predict outputs for inputs that has not yet been encountered.
  Example: minimize the length of the description of the hypothesis (Occam's razor).



Theory-Observation Circularity (contd.)

Empirical observation and acceptability of a hypothesis.

- The observation is framed in terms of a theory that also contains the hypothesis it is meant to verify or falsify.
- So, the observation can only arbitrate between the hypotheses within the context of the underlying theory.

Inductive bias.

- A modern notion used for learning algorithms.
- Assumptions that are used to predict outputs for inputs that has not yet been encountered.
  Example: minimize the length of the description of the hypothesis (Occam's razor).



Classification of the Sciences (Comte)

Mathematics, astronomy, physics, chemistry, biology and sociology.

- Classification to account for the diversity of the sciences without thereby losing sight of their unity.
- Classification is not meant to give a superficial unity but to avoid the fragmentation of knowledge.
- Movement from the general to the particular, and from the simple to the complex: moving from mathematics to sociology, generality decreases and complexity increases.
- Classification is the key to a theory of technology.
  - There exists a connection between complexity and modifiability: the more complex a phenomenon is, the more modifiable it is.
  - The order of nature is a modifiable order.
  - Human action takes place within the limits fixed by nature and consists in replacing the natural order by an artificial one.



Classification of the Sciences (Comte)

Mathematics, astronomy, physics, chemistry, biology and sociology.

- Classification to account for the diversity of the sciences without thereby losing sight of their unity.
- Classification is not meant to give a superficial unity but to avoid the fragmentation of knowledge.
- Movement from the general to the particular, and from the simple to the complex: moving from mathematics to sociology, generality decreases and complexity increases.
- Classification is the key to a theory of technology.
  - There exists a connection between complexity and modifiability: the more complex a phenomenon is, the more modifiable it is.
  - The order of nature is a modifiable order.
  - Human action takes place within the limits fixed by nature and consists in replacing the natural order by an artificial one.



Vienna Circle

A philosophical movement.

- Grew out of discussions at the Café Central in Vienna before World War I.
- Founding philosophers: Moritz Schlick, Otto Neurath and Rudolf Carnap.
- Included other mathematicians and philosophers such as Kurt Gödel.
- Propounded a philosophy of science called logical positivism.
  - A variant of positivist philosophy.
  - Influenced by the developments in the theory of formal deductive logic and mathematics.
  - The dominant philosophy of science in the first half of the 20th century.
  - Later criticism pointed out inherent inconsistencies.



Vienna Circle

A philosophical movement.

- Grew out of discussions at the Café Central in Vienna before World War I.
- Founding philosophers: Moritz Schlick, Otto Neurath and Rudolf Carnap.
- Included other mathematicians and philosophers such as Kurt Gödel.
- Propounded a philosophy of science called logical positivism.
  - A variant of positivist philosophy.
  - Influenced by the developments in the theory of formal deductive logic and mathematics.
  - The dominant philosophy of science in the first half of the 20th century.
  - Later criticism pointed out inherent inconsistencies.



Logical Positivism: Some Key Ideas

- All knowledge is based on logical inference from simple “protocol sentences” grounded in observable facts.
- Verificationism: a statement is “cognitively meaningful” only if there is method for determining whether it is true or false.
  - metaphysical, theological, and ethical statements do not sastisfy this criteria and cannot be called cognitively meaningful.
  - Such statements may have other kinds of meanings.
- Unified Science: development of a common language in which all scientific propositions can be expressed.
  - Reduction of terms of one special science to the terms of another.



Logical Positivism: Some Key Ideas

- All knowledge is based on logical inference from simple “protocol sentences” grounded in observable facts.
- Verificationism: a statement is “cognitively meaningful” only if there is method for determining whether it is true or false.
  - metaphysical, theological, and ethical statements do not sastisfy this criteria and cannot be called cognitively meaningful.
  - Such statements may have other kinds of meanings.
- Unified Science: development of a common language in which all scientific propositions can be expressed.
  - Reduction of terms of one special science to the terms of another.



Logical Positivism: Some Key Ideas

- All knowledge is based on logical inference from simple “protocol sentences” grounded in observable facts.
- Verificationism: a statement is “cognitively meaningful” only if there is method for determining whether it is true or false.
  - metaphysical, theological, and ethical statements do not sastisfy this criteria and cannot be called cognitively meaningful.
  - Such statements may have other kinds of meanings.
- Unified Science: development of a common language in which all scientific propositions can be expressed.
  - Reduction of terms of one special science to the terms of another.



Legacy of Different Positivist Movements

- Focus on science as a product, a linguistic or numerical set of statements.
- Concern with axiomatization:
  logical structure and coherence of these statements.
- Insistence on testability of statements:
  verification, confirmation, or falsification by empirical observations.
- Science is cumulative, transcultural and impersonal.
  - Results are dissociated from the personality and social position of the investigator.



Uncertainty and Induction: Examples

20 balls (B/W) in an urn;
a “random” sample of size 4 is drawn;
3 are black and 1 is white;
inference: 15 are black and 5 are white.

The proportion q of the *sample* has attribute A;
inference: the proportion q of the *population* has attribute A.

A proportion q of a population has attribute A;
An individual X is a member of the population;
inference: X has A with some probability (which depends on q).



Uncertainty and Induction: Examples

20 balls (B/W) in an urn;
a “random” sample of size 4 is drawn;
3 are black and 1 is white;
inference: 15 are black and 5 are white.

The proportion q of the *sample* has attribute A;
inference: the proportion q of the *population* has attribute A.

A proportion q of a population has attribute A;
An individual X is a member of the population;
inference: X has A with some probability (which depends on q).



Uncertainty and Induction: Examples

20 balls (B/W) in an urn;
a “random” sample of size 4 is drawn;
3 are black and 1 is white;
inference: 15 are black and 5 are white.

The proportion q of the *sample* has attribute A;
inference: the proportion q of the *population* has attribute A.

A proportion q of a population has attribute A;
An individual X is a member of the population;
inference: X has A with some probability (which depends on q).



Uncertainty and Induction: Examples

20 balls (B/W) in an urn;
a “random” sample of size 4 is drawn;
3 are black and 1 is white;
inference: 15 are black and 5 are white.

The proportion q of the *sample* has attribute A;
inference: the proportion q of the *population* has attribute A.

A proportion q of a population has attribute A;
An individual X is a member of the population;
inference: X has A with some probability (which depends on q).



Statistics and Induction

Statistical Inference/Induction (or Inferential Statistics).

- Methods for inferring from data that are subject to random variation.
  - Observational errors or sampling variation.
- Requirements of such methods.
  - Produce reasonable answers when applied to well-defined situations.
  - Applicable to a variety of situations.
- Possible conclusions of statistical inference methods.
  - An estimate.
  - A confidence interval.
  - Rejection of a hypothesis.
  - Clustering or classification of data points into groups.



Statistics and Induction

Statistical Inference/Induction (or Inferential Statistics).

- Methods for inferring from data that are subject to random variation.
  - Observational errors or sampling variation.
- Requirements of such methods.
  - Produce reasonable answers when applied to well-defined situations.
  - Applicable to a variety of situations.
- Possible conclusions of statistical inference methods.
  - An estimate.
  - A confidence interval.
  - Rejection of a hypothesis.
  - Clustering or classification of data points into groups.



Statistics and Induction

Statistical Inference/Induction (or Inferential Statistics).

- Methods for inferring from data that are subject to random variation.
  - Observational errors or sampling variation.
- Requirements of such methods.
  - Produce reasonable answers when applied to well-defined situations.
  - Applicable to a variety of situations.
- Possible conclusions of statistical inference methods.
  - An estimate.
  - A confidence interval.
  - Rejection of a hypothesis.
  - Clustering or classification of data points into groups.



Statistical Models

- Describes relationships between several **random variables**. Example: relationship between height and age of humans;

$$h_i = b_0 + b_1 \alpha_i + \epsilon_i .$$

  - h_i, α_i: height and age of the i-th person;
  - b_0: intercept; b_1: influence of age on height;
  - ϵ_i: error term determined by several factors including sex, race, etcetera.
- Models have to be tested against the data.
- More than one model may be proposed and comparison between models will be based on data.



Statistical Models

- Describes relationships between several **random variables**. Example: relationship between height and age of humans;

$$h_i = b_0 + b_1 \alpha_i + \epsilon_i .$$

  - h_i, α_i: height and age of the i-th person;
  - b_0: intercept; b_1: influence of age on height;
  - ϵ_i: error term determined by several factors including sex, race, etcetera.
- Models have to be tested against the data.
- More than one model may be proposed and comparison between models will be based on data.



Baye's Theorem and Its Influence

Let A be an effect which may arise from any of the mutually exclusive hypotheses $H_1, \ldots, H_k$. Then

$$\Pr[H_i|A] = \frac{\Pr[A|H_i] \times \Pr[H_i]}{\Pr[A]};$$

$$\Pr[A] = \sum_{i=1}^{k} \Pr[A|H_i]\Pr[H_i].$$

- View of probability:
  - Bayesian: to obtain the probability of a hypothesis, specify some prior probability and then update based on new data.
  - Classical (Laplace): ratio of favourable cases to all cases.
  - Frequentist: limit of relative frequency in a large number of trials.



Baye's Theorem and Its Influence

Let A be an effect which may arise from any of the mutually exclusive hypotheses $H_1, \ldots, H_k$. Then

$$\Pr[H_i|A] = \frac{\Pr[A|H_i] \times \Pr[H_i]}{\Pr[A]};$$

$$\Pr[A] = \sum_{i=1}^{k} \Pr[A|H_i]\Pr[H_i].$$

- View of probability:
  - Bayesian: to obtain the probability of a hypothesis, specify some prior probability and then update based on new data.
  - Classical (Laplace): ratio of favourable cases to all cases.
  - Frequentist: limit of relative frequency in a large number of trials.



Bayesianism and Induction

Keynes (1921), Carnap (1950): Inductive logic where Baye's theorem plays a central role.

- Specifies inductive support probabilities solely in terms of the syntactic structure of premise and conclusion sentences.
- Inductive probabilities in such a system are 'logical' in the sense that they depend on syntactic structure alone.
- Criticism: syntactic logical structure cannot be the sole determiner of the degree to which premises inductively support conclusions.
  - Applications to scientific contexts where the conclusion is a hypothesis, and the premises are observations.
  - The difficulty is that in any probabilistic logic that satisfies the usual axioms for probabilities, the inductive support for a hypothesis must depend in part on its prior probability.
  - No satisfactory general method to assign prior probabilities.



Bayesianism and Induction

Keynes (1921), Carnap (1950): Inductive logic where Baye's theorem plays a central role.

- Specifies inductive support probabilities solely in terms of the syntactic structure of premise and conclusion sentences.
- Inductive probabilities in such a system are 'logical' in the sense that they depend on syntactic structure alone.
- Criticism: syntactic logical structure cannot be the sole determiner of the degree to which premises inductively support conclusions.
  - Applications to scientific contexts where the conclusion is a hypothesis, and the premises are observations.
  - The difficulty is that in any probabilistic logic that satisfies the usual axioms for probabilities, the inductive support for a hypothesis must depend in part on its prior probability.
  - No satisfactory general method to assign prior probabilities.



Bayesian Inference

- Refers to the use of a prior probability over hypotheses to determine the likelihood of a particular hypothesis given some observed evidence.
- Consistency of a given hypothesis with collected evidence.
- With growing evidence, the confidence in a hypothesis changes; with enough evidence, the degree of confidence should become either very high or very low.
- Results will be biased based on a-priori assumptions.



Universal Method of Induction

"Statistics is the universal tool of inductive inference, research in natural and social sciences, and technological applications.

Statistics, therefore, must always have purpose, either in the pursuit of knowledge or in the promotion of human welfare."

– Prasanta Chandra Mahalanobis
(2nd December, 1956)

