BEYOND THE TRADITIONAL FRONTIERS OF STATISTICAL SCIENCES: A CHALLENGE FOR THE NEXT DECADE*

by

Pranab K. Sen

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 1861

March 1989

BEYOND THE TRADITIONAL FRONTIERS OF STATISTICAL SCIENCES : A CHALLENGE FOR THE NEXT DECADE *

PRANAB KUMAR SEN
University of North Carolina,
Chapel Hill, NC 27599,USA

Statistical sciences, albeit of relatively recent origin (compared to the other basic sciences), have made a remarkable stride during the past few decades. Most of these developments have taken place in the sector of basic theory and methodology pertinent to the traditional areas of applications, such as the agricultural sciences, biometry, medical studies (clinical trials), physical sciences, reliability theory and quality control, demography, econometrics, social sciences and few other related fields. Statistical quantum physics paved the way for a sound foundation of some relevant probaility theory and stochastic processes with a variety of fruitful applications. Electrical engineering, industrial engineering, operations research and systems analysis all have added new dimensions to statistical methodology. All of these are on a sound footing and deserve special commendation on an unequivocal basis.

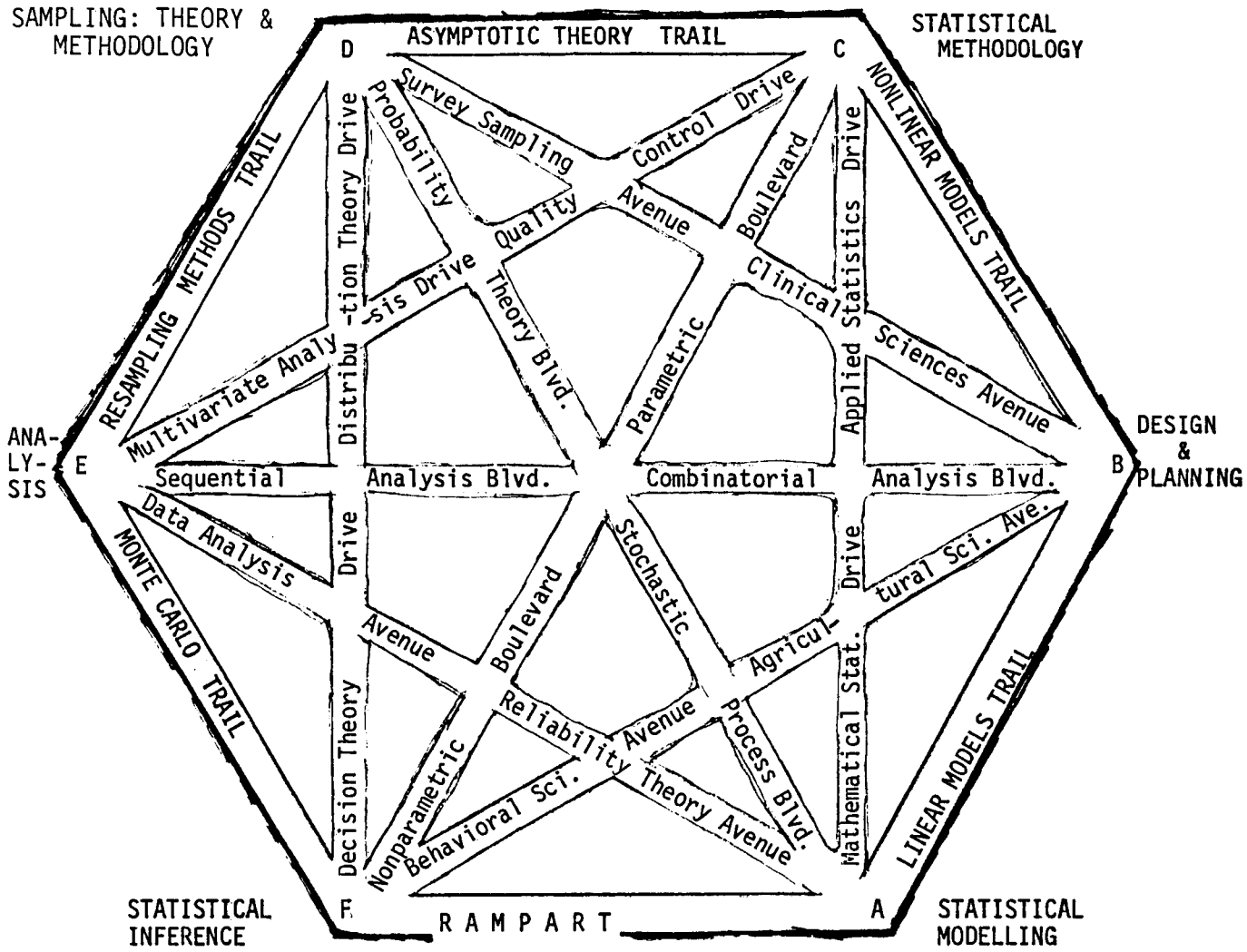
The past two decades have witnessed a phenomenal growth of research methodology in a variety of experimental, epidemiological and environmental sciences where not only statistical priciples occupy a focal position but also there is a genuine need for developing more adoptable statistical theory and methodology in these novel, non-standard and diverse setups. Statistical modelling plays a fundamental role in this context. The main objective of the current study is to stress some of these aspects of statistical sciences and to focus on the challenges arising in this context.

In order to pose some of these challenging problems in a proper perspective, perhaps, it will be quite appropriate to have a critical appraisal of our current standing and then to focus on some of these new areas of applications where the challenges are to be met. Afterall, statistical sciences have the genesis in real applications, and the strive for success also rests on the scope for accommodation of new areas of fruitful applcations. As such, we proceed by examining first the traditional frontiers of statistical sciences along with the major developments within this territory. A composite picture for this is presented in the following:

---

* Platinum Jubilee Lecture in Statistics at the 76th Session of the Indian Science Congress Association at Madurai, S.India, on January 10, 1989.

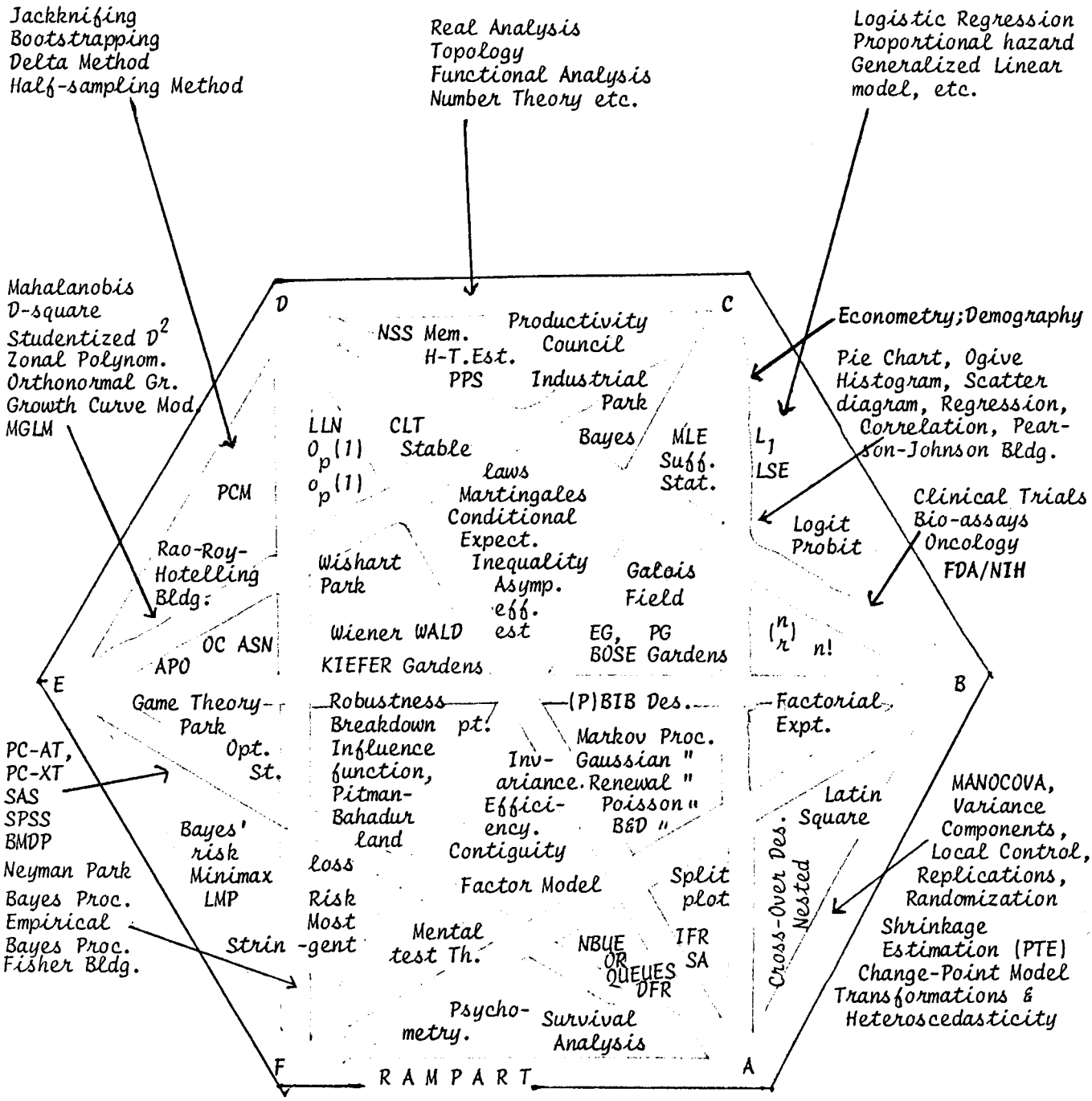
FORTRESS STATISTICAL SCIENCES : 1989



As has been noted earlier that the genesis of statistical sciences dates back to the basic need for quantifications in other basic sciences. It is therefore of no surprise that a variety of nomenclatures pertains to statistical sciences. The most popular ones are listed below :

BIOSTATISTICS/BIOMETRY
APPLIED STATISTICS/ Quality Control
INDUSTRIAL ENGINEERING/Operations Research
SYSTEMS ANALYSIS
ECONOMIC STATISTICS/ Management Sciences
EPIDEMIOLOGICAL STATISTICS
EXPERIMENTAL STATISTICS
AGRICULTURAL STATISTICS
EUGENICS
POPULATION STUDIES
E C O L O G Y
MATHEMATICAL STATISTICS
STATISTICS
MEDICINAL STATISTICS
APPLIED PROBABILITY
DATA MANAGEMENT
ENVIRONMENTAL STATISTICS
ACTUARIAL STATISTICS

To complete the (residential) picture of the Fortress, let us superimpose some of the major developments in these areas. Towards this, we have the following



Each of the six major sectors in the fortress plays a basic role, and their appropriateness depends on the specific situations at hand. It is quite conceivable that more than one of these sectors are functionally operative in any particular problem arising in any scientific study. To stress further the basic role of these sectors, let us consider an (alphabetical) listing of the basic and applied sciences where statistical principles are indispensible. This list is by no means complete. Rather, it portrays a comprehensive picture of the vast domain of applicability of statistical reasoning in every walks of life.

A. ANTHROPOMETRY , AGRICULTURAL SCIENCES, ARTIFICIAL INTELLIGENCE
B. BIOLOGICAL SCIENCES, BIOMETRY
C. COMPUTER SCINCE, CLINICAL TRIALS, CRIMINOLOGY
D. DEMOGRAPHY, DATA COLLECTION AND ANALYSIS, DAIRY SCIENCES, DRUGS
E. ECOLOGY, EPIDEMIOLOGY, ENVIRONMENTAL SCIENCES, EPILEPSY, ECONOMICS, ENERGY
F. FAMILY PLANNING, FOLK MEDICINE, FORESTRY, FOOD TECHNOLOGY
G. GAMBLING, GENETICS, GEOGRAPHY, GEOSCIENCES, GENETIC TOXICITY (MUTAGENICITY)
H. HEALTH SCIENCES, HYGIENE, HYDROLOGY, HUMANITY
I. IMMUNOLOGY (AIDS), IMAGE PROCESSING(DNA), INDUSTRIAL RESEARCH
J. JURISPRUDENCE , JUVENILITY
K. KINEMATICS, KINETICS
L. LEPROSY, LINGUISTICS, LANDSCAPING, LEGISLATURE
M. MANAGEMENT SC., MARINE SC., MATHEMATICS, MEDICAL SC., METEOROLOGY
N. NATURAL SCIENCES, NEURO-BIOLOGY, NUTRITION
O. ONCOLOGY, OPERATIONS RESEARCH, OCEANOGRAPHY, OCCUPATIONAL HAZARDS
P. PHYSICAL SCIENCES, PUBLIC HEALTH, PSYCHOLOGY
Q. QUALITY CONTROL, QUANTITATIVE SCINCES, QUARRYING, QUARANTINE
R. RADIATION EFFECTS, RELIABILITY (SYSTEMS ANALYSIS)
S. SOCIAL SCIENCES, SEMIOLOGY, SEXOLOGY, SURGERY
T. TOXICOLOGY, TELEPATHY, TRAFFIC ENGINEERING, THROMBOSIS
U. UNCERTAINTY,UNEMPLOYMENT, UNESCO, UNIVERSE
V. VETERINARY SCIENCE,VENEREAL DISEASES, VIROLOGY
W. WAR SCIENCES, WATER , WRESTLING
X. XEROGRAPHY, XENODERMA
Y. YOGA, YACHTING
Z. ZOOGEOGRAPHY, ZOOLOGY, ZYMOLOGY

Since space and time are a premium, let me pick up a handful of these areas and illustrate some of them in some detail. Towards this, we have the following:

A : AIR, AQUA (Water), ARTERIOSCLEROSIS , ARTIFICIAL INTELLIGENCE

E : ENERGY, ECOLOGY, ENVIRONMENT

I : IMMUNOLOGY(AIDS), IMAGE PROCESSING(DNA)

O : ONCOLOGY , OCCUPATIONAL HAZARDS

U : UNCERTAINTY , UNIVERSE(Ozone / Greening Effect)

We start with a glossary of ENVIROMENTAL Problems which encompass a greater amount of these areas. Since most of these problems have become quite noticable in the current time, we avoid any further description of the actual details.

I. ENVIRONMENTAL POLLUTION:

Taj Mahal, Agra, India Venice, Italy	AIR-BORNE Sulphur Deposition : OIL Refineries
Love Canal, USA Bombay Marina, India New Jersey , USA Berlin (East) Waterfront, Northern W.Genmany Rivers	CHEMICAL DUMPINGS : Water Contaminations
Chernobyl, USSR Three Miles Island, USA Savanah River Nuclear Plant, S.Carolina, USA	WATER CONTAMINATIONS : Radiation Effects; AIR CONTAMINATIONS : Radiation Effects.
Hiroshima & Nagasaki, Japan	RADTIATION CASUALTY : Nuclear Bombs
Nuclear Power Plants' Waste Deposites Sites: All Over the Globe	Water Contaminations (Subsoil) due to Radiation
Thermal Power Plants in India and Elsewhere	AIR-BORNE CARBON Deposition: Black Lungs !
Oil-Slick : Coastline of France, California (USA),...	Ecological Disaster of Marine Biology
Chemical Warfare in Iraq- Iran War , Agent ORANGE in Vietnam War	Spontaneous Casulty and long term Hazards
Bhopal Episode (Union Carbide) and others	Catastrophic Mortality & Morbidity due to Chemical actions.
Industrialized Nations	Greening Effect : OZONE Layer Problem
Deforestation: Natural or Deliberate Actions	Cooling Effects / Erosion
Automobiles/Planes/ Factories.....	Smog & Respiratory Diseaes.
Ganges, Varanashi, India	Acute Water Pollution due to dumping of Human Waste and partly cremated cadavers.
Acid-rain	Atmospheric Pollution Recycles.
Solid Waste	Landfill Problem in any urban area.
RADON in the Basement	It is no longer at Random

Instead of going on with an uncountable number of such major problems, let me assert the statistical issues relating to some specific ones. Although the other problems can be treated in a similar manner, there may be specific needs for some modifications to suit the actual experimental setups.

Solid Waste Problem : The Garbage Dilemma.

It is a major problem for every urban area in every industrialized nation ! The dimensions are no less for big cities like Bombay, Calcutta or even Madurai !

Against the traditional procedure of LANDFILLing (or dumping), there are some other alternatives :

(i) RECYCLE a greater part of the Waste,

(ii) LANDFILL the remainder in modern engineered safe sites, and

(iii) Use a RESOURCE Recovery Technology to reduce non-recycled waste to some extent.

The conventional dumping of garbage into Sea or Underdeveloped Landsites is Hazardous in many respects : Toxic material in the garbage may produce far reaching hazardous effects, cause serious pollution problems and contaminations in undesirable manners. A safe and proper Landfill may require a generous amount of land and many safety precautions, and may therefore be hard to locate. Newspapers, card boards, aluminum cans, glass bottles and jars and some other materials in the garbage, if properly sorted at the collection stage, may be recycled upto 50% . Massburn Resource Recovery Plants may reduce the landfill disposal volume of non-recycled solid waste by about 80%. However, the air pollution impacts are needed to be monitored closely. Composting relates to a natural conversion of the organic portion of solid waste to humus by micro-organism activity. However, in this process, removal of hazardous and non-biodegradable materials is required. Refuse Derived Fuel : Garbage into Fuel, can be used as supplemental Fuel with about 50% heating value of coal and having low sulpher content.

It is a complex problem, and plausible recommendations include the following:

i) We must preserve the capacity of existing landfills and begin immediately to acquire and permit new facilities or reserve sites for future use.

ii) Need to develop long-range strategies to significantly reduce reliance on conventional landfills

iii) A comprehensive Recycling plan needs to be developed, preferably on a regional basis

iv) Composting plants are desirable on a small area basis : Regional plants may not be feasible.

v) Regional Refuse Derived Fuel plants are preferable

vi) Public education and statistical evaluations require utmost attention.

Any master plan in this context must take into account (i) epidemiological factors, (ii) environmental issues and (iii) awareness to recycle and other alternatives. It is clear that STATISTICAL MODELLING plays a vital role in this respect, and the standard setup of simpler statistical models may not be that appropriate here. The anatomy of garbage may have a lot of geographical variation and similar considerations dominate the adoptibility of alternative means : Statistical considerations are thus very important

Acid Rain : Genesis, Appraisal and Statistical Considerations :

In the Indian subcontinent, from the age of Kalidas to Rabindra Nath, the RAIN has been the messenger , the lyric and the reliever :

Let the peacock in thy heart dance with the tune of monsoon,
and let thy emotion in a riot of cheerful colors be in full bloom.
Let the rivers flood their banks in sheer generosity,
and the golden harvest echo thy thanks in total entity!
But,in an era of immensurable atmospheric pollution,
can monsoon bestow you with any hygienic resolution?
With the seeded clouds, what are we having -
lots of pollutants sans thy warm blessing !
Rain drops soaked with sulpher, acids and smogs,
city streets flooded with human waste and garbage in tons!
Normal living utterly disrupted with dreadful diseases,
from eyes and lungs to the intestine, to the aging processes.

What are we having now in the rain :

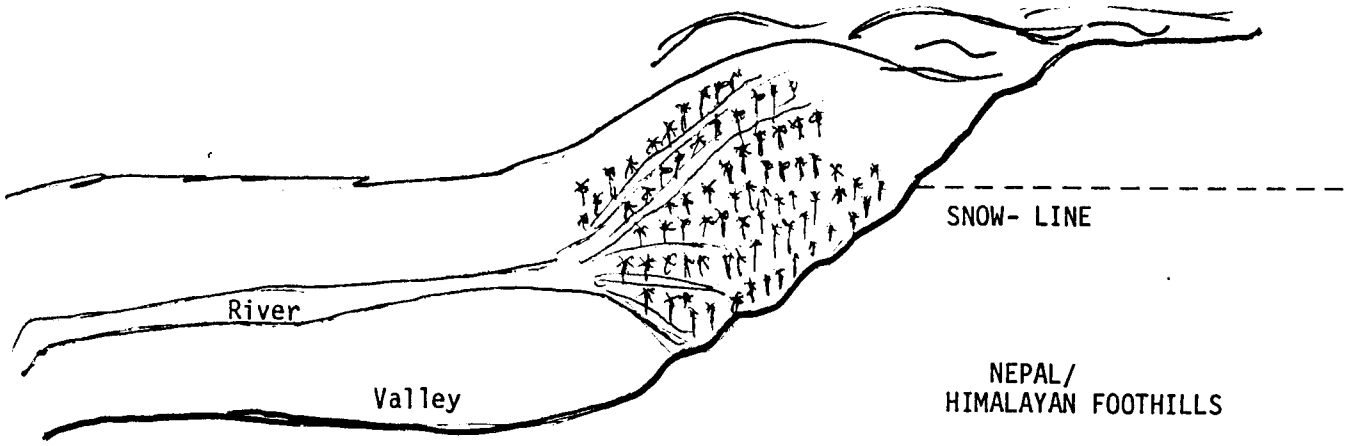
Carrier of environmental pollutions,
but, whom to blame for our own creations ?

Combustion of fuel and gasoline, thermal and nuclear power plants emissions, absorption of ultra-violet rays (due to the thinning of the Ozone layer), and what not ? All of these factors account for the acid rain phenomenon. In the industrialized western countries , the snow in the winter has a similar effect: it leads to the absorption of toxic materials in the soil and contaminates the water resources too !

While the environmental scientists can identify plausible sources of pollutions and explain their impacts in terms of physical effects, the problem is much more complex, and a statistician (with an indepth knowledge in the environmental sciences) can not only identify the demographic and epidemiologic aspects much more clearly, but also suggest suitable models for such a phenomenon to be describable in a simple and estimable quantitative form . We have a long way to go along this path!

Mother Earth : Vegetation Baldness (Vanishing Tree lines) :

Although we are lamenting over the natural calamities occurring in our highlands and the planes, have we paid due attention to the conservation of plants and ecology in this planet ? Deforestation not only affect the living style in the foothills, it has serious impacts on the flattening of the riverbeds due to erosion at the foothills and also contribute to the disasterous flooding in the planes. As a largely over-simplified version of this complex problem, we may turn towards Nepal or other territories in the Himalayan foothills where ecological disasters are quite imminent!



DEFORESTATION: (1) Agricultural Needs
(2) Energy : Cooking & Heating Purposes
(3) Forest-fire and Avalanches

A) Less Vegitation and less Absorption of CO_2
B) More EROSION due to RAIN/SNOW/AVALANCHES
C) LANDSLIDES resulting in less Fertile Lands for Agricultural Utilizations.

NEED FOR AFFORESTATION:

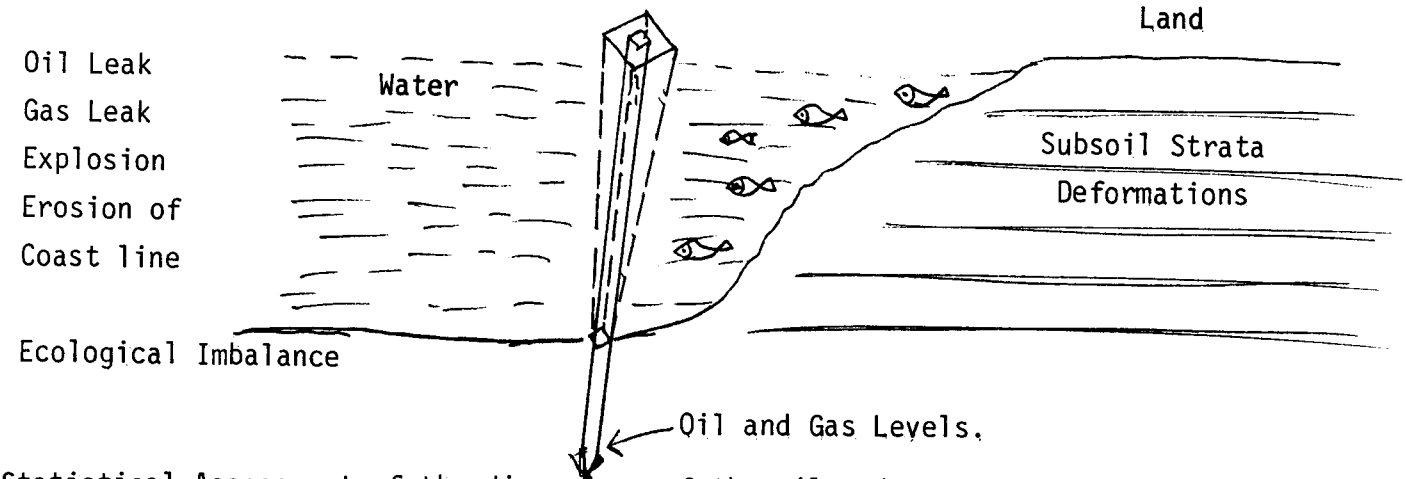
(1) Demographic Considerations
(2) Ecological Considerations
(3) Exploration of Alternative Sources of Energy

Statistical Considerations are important in each Sector:

Statistical Modelling, Planning and Designs dominate the scenario

INDUSTRIALIZATION AND POPULATION EXPLOSION have taken over HARDWAR-RISHIKESH, all the way to the PRAYAGs . Our DEFENSE Network has been extended to the northern Frontiers all the way to the Tibbet boarder. What about our ECOLOGICAL and ENVIRONMENTAL PLANNINGS and UNDERTAKINGS ? STATISTICIANS BE ON ALERT !

II. OFFSHORE OIL EXPLORATION : MARINE BIOLOGY & ECOLOGICAL IMPACTS

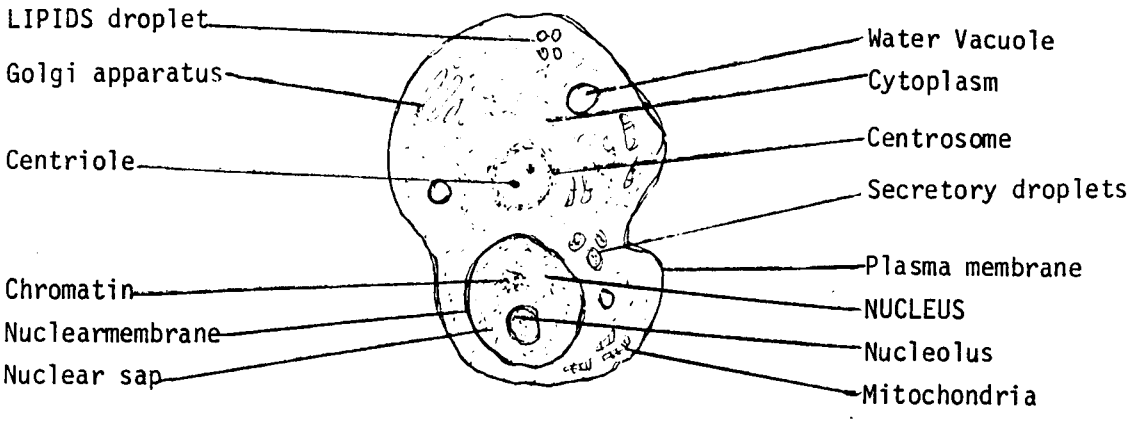


Statistical Assessment of the dimensions of the oil and gas reserves !
Detailed study of the Marine biology and the ecological considerations !
Statistical risk analysis of Oil/gas leaks, explosions and other forms of disasters !
Deformation of coastline and sub-soil strata due to energy extraction !

III. GENOTOXICITY : ENVIRONMENTAL IMPACTS.

To start with we consider the basic GENETIC setup in human beings.

THE CELL



Chromatin: The portion of the nucleus of a cell that is readily stained by dyes. It forms a fibrillar network of granules of various shapes and sizes, composed of DNA combined with a basic protein and is the carrier of the GENE.

Chromosome: A body in the cell nucleus that is the bearer of the GENE. It has the form of a delicate Chromatin filament during interphase, contracts to form a compact cylinder segmented into two arms by the centrosome during metaphase and anaphase stages of cell division, and is capable of reproducing its physical and chemical structures through successive cell divisions.

DNA(Deoxyribonucleic acid): Type of nucleic acid containing deoxyribose as the Sugar component and found principally in the nuclei (Chromatin/Chromosomes) of living cells, usually loosely bound to Proteins. Considered to be the autoreproducing component of Chromosomes and of many VIRUSES and the Repository of Heriditary characteristics.

RNA(Ribonucleic acid): Found in all cells, in both nuclei and cytoplasm, and in particulate/nonparticulate form, also in many Viruses.

GENE: A complex, self-reproducing molecule that constitutes the unit of heridity. As a rule, a full complement of genes is present in every (somatic /germ) cell of the body, arranged in linear order, like beads on a string, at definite points (or loci) in chromosomes. There are 23 pairs of chromosomes in humans and since these chromosomes occur in pairs, the genes are also paired. There are some thousands of different kinds of genes in human beings. For any cell, the genes govern the functions of the cell and are primarily responsible for its propoerties and behavior. There are literally thousands of traits (both qualitative and quantitative) and genes are associated with these.

GENETIC INFORMATION is contained in the relatively simple CHAIN MOLECULE that constitute DNA and RNA. Each of these chains consists of four different kinds of units called NUCLEOTIDES, so connected that thousands of them can join chemically end to end

to form a long string. Each nucleotide consists of a sugar molecule (Ribose in RNA and Deoxyribose in DNA) , a phosphate group and a nitrogeneous base. In DNA four nitrogeneous bases are involved: the purines adenine and guanine, and the pyrimidines thymine and cytosine; RNA the pyrimidine base uracil replaces thymine.

Watson and Crick, and later on, Khorana and his associates have all developed some THEORETICAL MODELS for the DNA Structure which account well for the four essential properties of the Gene : Specificity, Replication, Mutation and Function. The Genes usually produce (through MUTOSIS or Cell Division) EXACT COPIES of themselves for hundreds of Cell Generations. But, sometimes a sudden and permanent gene change takes place - this is termed a MUTATION.

The genes of a cell thus carry a set of messages. They give the cells Chemical Instuructions about how to manufacture certain kinds of molecules. These instructions are passed from the DNA to the RNA. RNA transfers the instructions out of the Nucleus to certain structures in the cell called Ribosomes - and the ribosomes determine the way PROTEIN MOLECULES are formed. Thus the DNA instruct the arrangements of Amino Acids. A small change in a DNA message can cause one amino acid to replace another - results in a different kind of protein, and this in turn can bring a change in the structure.

VIRUS : Genetic material in a Virus particle is constituted in two closely paired helical chains, coiled around a common axis- each chain consists of four kinds of nucleotides.

Virology in the light of DNA/RNA ! How about CANCER ?

AIDS (Acquired Immuno Deficiency Syndrome): HIV (Human Immunodeficiency Virus)
Attacks the White Blood Cells and paralyses the Immunity.
Clinical Aspects (Perhaps Genetical Mutations)
Epidemiological Aspects
Environmental Aspects.

BIOTECHNOLOGY : Control of Human Traits Through Forced MUTATION !
Ethics & Risks !

GENOTOXICITY : To What Extent Can it be Avoided ?

What is our Contribution in this Complex Area:

(i) Image Processing : An Immensurable Dimensional Data Analysis !
(ii) Statistical Modelling : Goodness of Fit Dominates the Scenario!
(iii) Drawing of Statistical Conclusions : Beyond the Neyman-Pearson Domain!
(iv) Unconditional Surrendering to Computer Graphics and Programmings ?

This is by far one of the most challenging areas for statisticians to contribute to the research methodology. Statistical modelling is the most hardest part of this task. A model must be flexible enough to allow standard statistical tools to be operative, and at the same time, it should be practical enough to allow a natural physical interpretation of the phenomenon in the light of the experimental factors and statistical variations.

" The presence of TOXIC CHEMICALS in our Environment is one of the grimmest discoveries of the industrial era. Rather than coping with these hazards after they have escaped into our environment, our primary objective must be to prevent them from entering in the first place." [President Carter, 1977]

A MUTAGEN is an agent, e.g., X-rays or a chemical, that is capable of effecting heritable change in genetic material, be it at one specific DNA base-pair or at the Chromosomal level.

(i) Mutagens may exert their effects on Germ Cells,thereby altering the human gene pool and possibly inducing genetic defects in future generations

(ii) Mutagens that attack Somatic Cells may lead to Heart Disease, Aging and Developmental Birth Defects.

(iii) Mutagens play a significant role in the CARCINOGENESIS Process.

MUTATION : A genetically based change, generally expressed as a change in drug resistence, a new metabolic requirement ,or an altered gene product.

BASE SUBSTITUTION: A change in the genetic 4-base code involving the replacement of one DNA base for another.

FRAME SHIFT: Insertion or deletion of bases such that there is an offset of the triplet reading frame of DNA, resulting in a string of incorrect Amino Acids as the gene is being translated into protein.

REARRANGEMENT: Deletion, insertion, inversion or translocation of DNA sequences.

CHROMOSOME ABERRATION: Microscopically visible change in chromosome structure.

ANEUPLOIDY: Change in chromosome number .

SISTER CHROMATID EXCHANGE : An easily scored exchange of chromosome strands within replicating chromosomes.

MICRONUCLEI: Small secondary nuclei within cells, indicating breakage of chromosomes.

DNA REPAIR: Evidence that one or more types of repair of DNA damage has taken place.

DNA ADDUCT: A binding of exogenous chemical to DNA

DNA STRUCTURE DAMAGE: Molecular weight change, strand breakage or other gross alterations in the structure of the DNA

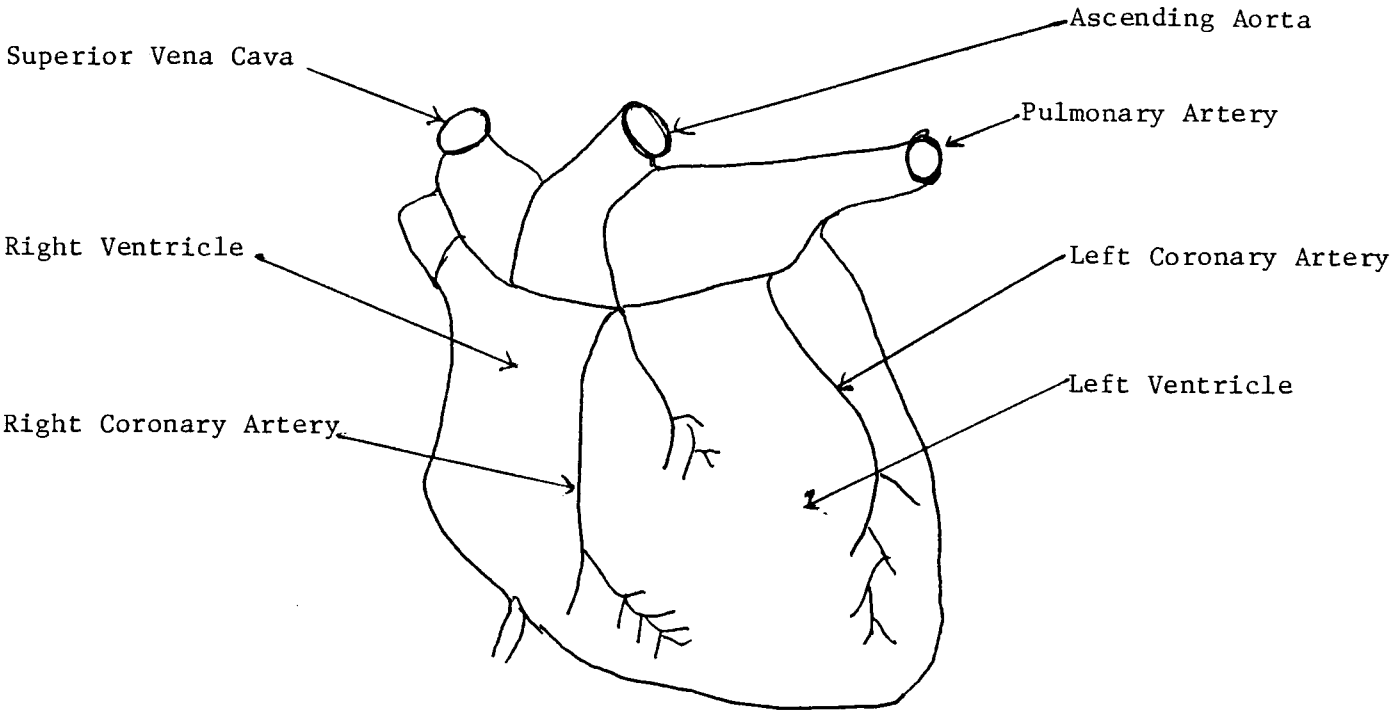
[*Tests for these are referred to as GENOTOXICITY TESTS* .]

MARKERS of Genotoxicity are going through an evolutionary change with a matching support from the BIOTECHNOLOGICAL Developments. Some markers of genotoxicity are already being incorporated into epidemiological studies. There is a vast amount of opportunities for more sophisticated scientific studies in this direction. It should come as no SURPRISE that the process of VALIDATION and then Application to Epidemiology will present many interesting and important STATISTICAL CHALLENGES.

Fortunately, public awareness of such GENOTOXICITIES has become quite noticable in the recent past, and given this factor, further developments are bound to occur in near future. Statisticians have a great role to play in this context too.

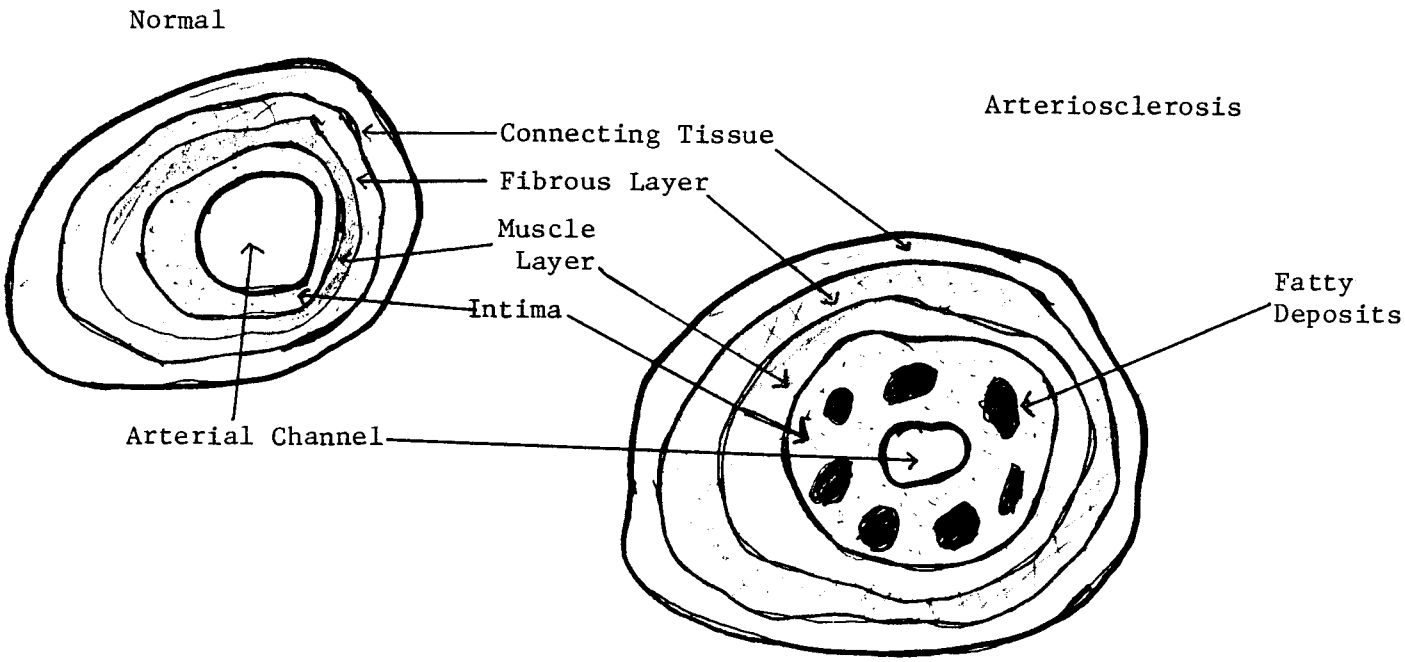
IV. ARTERIOSCLEROSIS.

We start with the human heart:



The arteries may become thick and hard and lose their supple elastic quality: Normal Aging Process. But Hypertension, Overweight, Smoking, Diabetes, Inactive (P.E.) all may contribute significantly towards the increase of CHOLESTEROL in the Blood.

Efffects : Angina Pectors/ Heart Attack,
Strokes (Ascend. Aorta) , Perkinson's disease etc.



Is it a normal aging process ?

Is it due to some ascribable factors ?

ATHEROSCLEROSIS is the process in which this disorder takes place, and
ARTERIOSCLOROSIS is the terminology for this disorder.
CHOLESTEROL has been identified as the most significant factor in this context.
LDL (Low Density Lipoproteins) are the bad elements , while
HDL (High Density Lipoproteins) are supposed to be beneficial.
Risk Ratio is defined as (LDL + HDL)/(HDL)
American Medical Association Rule of Thumb is that this risk ratio should be ≤ 3.5.
Also,

	LDL + HDL measure	$\leq$	200	: Normal level
200 <	LDL + HDL measure	<	240	: Border line High
	LDL + HDL measure	$\geq$	240	: High.

Whole Milk, Animal Fats (Crabs & Shrimps included), Butter, Coconut Oil, Palm Oil, Egg, Cheese, Beef, Pork etc., all are supposed to be liberal contributors for the increase of LDL.
Moderate amount of Physical Exercise, Low Animal Protein Diet (e.g., FISH), High Fibre/Carbohydrate Diet, Less Tension and Regular Habits are supposed to be beneficial.
Recent Clinical Trials have all supported that

Lowering of LDL and increasing of HDL improve the picture considerably !

There are potentially scores of statistical issues connected with this Arteriosclerosis problem . To mention just a few important ones, we consider the following :

(i) What Response Variables would you choose in a quantitative assessment of arteriosclerosis ?

(ii) If one has to use arterial cross sections, what are the most strategic points in this respect ?

(iii) To what extent modern bio-technology can be adapted to improve the recording of observations and increasing the sensitivity of the statistical procedures ?

(iv) In view of (iii), what should be an overall Objective of the study ?

(v) Is there a better way to define the risk ratio and/ or the classification of the state according to the cholesterol level ?

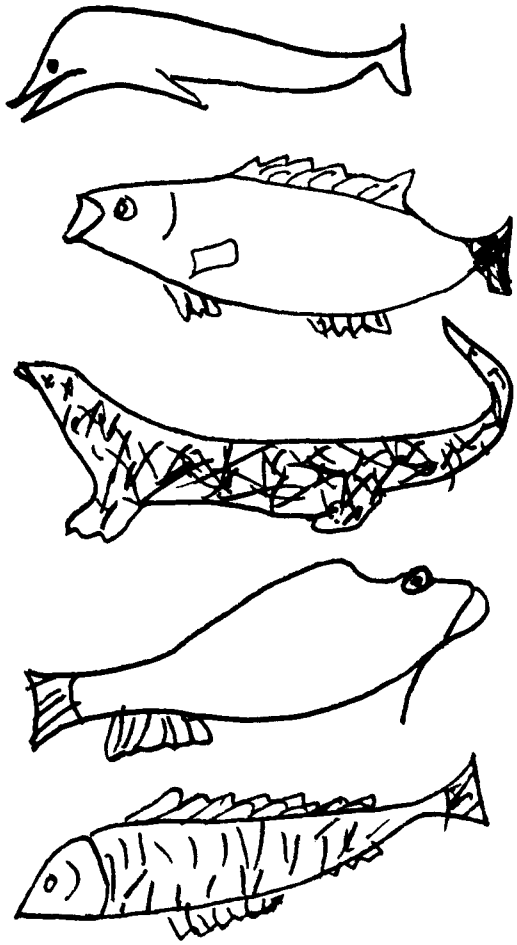
(vi) With respect to the classifications of diet affecting cholesterol, what is a sound statistical design to draw statistical conclusions in a valid and efficient manner ?

(vii) Since cross sectional study of arteries are made following the termination of life, there is a greater need to develop a statistical plan on a sound epidemiological and environmental basis. To what extent Accelarated Life Testing Techniques are appropriate in such a study ?

(viii) To what extent standard statistical methodolgy can be used in this study? The usual assumptions are not found to be appropriate in the current context.

V. STATISTICS & ECOLOGY

We consider as an illustration a simple Eco-Cycle :



Whale / Dolphin

Shark / Raghava Boal

Crocodile / Alligator

Scolpin / Tuna

Salmon/ Rock / Trout etc.,

Mixed Water Fish : Shrimp, Crabs etc.

Sweet Water Fish : Croker, Bass, Catfish etc.

There is an intricate network of Ecology : Pulling any string may affect the whole ecological system :

EX. WHALE Hunting in the Pacific and Atlantic Oceans.

(i) Likely to lead to the extinction of a species

(ii) Disasterous effects on Marine Biology/ Marine Geography

(iii) Ecological imbalance !

The ecological chain of the marine animals is indeed very complex. Not only marine biology and marine geography are pertinent in this context, but also there is a genuine need to study the environmental impacts in a critical way. Statistical Modelling is thus a very important task in this respect. Sampling techniques may also be quite different from the conventional ones, and need careful attention. Weighted or length-biased sampling methods have been considered by some workers, and more practical ones are needed to be developed.

I could continue in this fashion in citing other fruitful areas where statistical methodology plays a vital role . However, I would rather check that temptation, and try to summarize the observations in the form of the following comments :

I. In agricultural sciences / biometry/ clinical trials/ physical sciences, generally, an experimentation involves a comparative study of a **control** and one or more **treatment** or experimental groups. In most of the problems referred to above, this conventional setting may not be that appropriate.

II. In the conventional setup, an experiment leads to a well specified model involving a set of unknown parameters and some form of the statistical variations, and the problem is to draw statistical conclusions (subject to desired level of precisions) from the experimental data set in an efficient manner. In the current context, the formulation of a suitable model is the first and foremost task. A model needs to be very relevant to the actual experimental setup, must take into account the intricacies ascribable to the various assignable factors,and at the sametime, must pay sufficient attention to the mechanism of chance variations associated with the experimental setup. In this context, we may note that Nature plays a vital role in such environmental and epidemiological studies, and that human being may not have much control over Nature. Thus, the list of plausible assignable factors may not be generally very tenuous, nor the actual mechanism of chance variations be ascribable in very simple terms.

III. Given (I) and (II), the challenging task is to develop appropriate statistical models in such diverse problems. Moreover, in the majority of cases, an adequate statistical model may involve a large number of parameters. With the increase of the number of parameters, the precision of statistical conclusions may go down rather drastically, unless the number of observations is made to increase adequately. This may not always be possible. In such a case, it may be necessary to switch to a reduced model with lesser number of parameters. Such reduced models are to be considered with utmost care and attention to the actual experimental setup. There are numerous examples where such reduced models were totally inappropriate.

IV. Often, in the problems mentioned above, the sampling design is quite different from the conventional ones. This in turn lead to a setup where the conventional assumption of independent and identically distributed random variables (relating to the observations) may not stand valid. Thus, the conventional statistical methodology may not work out properly in such problems. Hence, contingent on the formulation of suitable statistical models and sampling designs, there is a need for the development of novel statistical methodology to handle such problems in a more effective manner.

V. In drawing statistical inference in a conventional setup, the Neyman-Pearsonian approach, Bayes methods and the Wald decision theoretic approach dominate the scenario. In most of the problems referred to above, the risk functions are generally of much more complex form. Also, the cost-benefit formulation or some other formulation taking

into account the ethics and experimental restraints seems to be more appropriate. Thus utility-functions oriented formulations may appear to be better in this context. There is thus a need to extend the conventional statistical analysis methodology to such utility-function oriented setups.

VI. In physical sciences (viz., molecular movement), statistical mechanics may as well be described in terms of appropriate stochastic differential equations where in many cases some physical interpretations may also be given to the algebraic constants appearing in these equations. It may not be unnatural to think of such stochastic differential equations in the formulation of statistical models appropriate to the problems referred to above. However, providing physical interpretations to the algebraic constants in such equations may be a harder problem!

There are other problems associated with the formulation of statistical methodology for the problems referred to above. We may be able to resolve these problems more satisfactorily if we choose to work closely with the environmental, epidemiological and ecological scientists, understand the basic principles, formulate the problems in a manner compatible with the experimental setup, and try to develop adequate and novel methodology. This prospect also depends on the reciprocity of the other scientists in these disciplines in their willingness to adopt more statistical methodology in a more sophisticated manner. Only then, we will be able to meet the challanges.

A little over sixty years ago, R.A. Fisher in his excellent treatise : Statistical Methods for Research Workers [Oliver & Boyd, Edinburgh, 1925] aroused the scientists in a variety of disciplines of the potential impact of statistical methods in their research. Soon after, Professor P.C. Mahalanobis initiated a highly effective research methodology in crop forecasting and related areas. Half a centruy later, although these methodologies retain their effectiveness, we encounter a greater task of updating them with a view to meeting the demands of the new avenues of sciences mentioned here. Statistical modelling seems to be the most important aspect of this challenge, and contingent on that other developments should follow. We look forward to this.