ISSN: 0975-833X

Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 12, Issue, 01, pp.9616-9620, January, 2020

DOI: https://doi.org/10.24941/ijcr.37121.01.2020

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

RESEARCH ARTICLE

USE OF MATHEMATICAL STATISTICS IN SOCIAL SCIENCES

***Dr. Deepak Kumar Routray**

ISPAT Autonomous College, India

ARTICLE INFO

Article History:
Received 12th October, 2019
Received in revised form
28th November, 2019
Accepted 09th December, 2019
Published online 30th January, 2020

Key Words:

Experiments, Hypotheses, Mathematical Methods, Methodology, Social Sciences, Statistical Inference, Significance Testing.

ABSTRACT

Over the years, mathematical statistics have become increasingly important in the social sciences. In fact, the use of mathematical statistics methods is now ubiquitous: Almost all social sciences rely on statistical methods to analyze data and to form hypotheses, and almost all of them use more or less a range of mathematical methods to help us understand the social world to some extent. Focusing on this the paper analyse and surveyed a variety of mathematical methods that are used in the social sciences and argued that such techniques, in spite of several methodological objections, can add extra value to social scientific research. It also discus some of their philosophical questions and focused on methodological issues in statistics- the part of mathematics that is most frequently used in the social sciences, in particular in the design and interpretation of experiments. The paper also analyzes the emergence of the rationale behind the ubiquitous significance tests, as well as explained the pitfalls to which many researches fall prey when using them. Finally, after comparing significance testing to rivalling schools of statistical inference, the recent trends was discussed in the methodology of the social sciences.

Citation: ***Dr. Deepak Kumar Routray.*** **2020.** "Use of Mathematical Statistics in Social Sciences", *International Journal of Current Research*, 12, (01), 9616-9620.

INTRODUCTION

Over the last few decade mathematical statistics have become increasingly important in the social sciences[1].A look at the history quickly confirms this claim. At the beginning of the 21st century most theories in the social sciences were formulated in qualitative terms while quantitative methods did not play a substantial role in the formulation and establishment of them moreover, many practitioners considered statistical methods to be in appropriate and simply not suited to foster our understanding of the social domain. Mathematical statistics methods were basically used and it became relatively uncontested that they were of much value in the social sciences. In fact, the use of mathematical statistics methods is now ubiquitous: Almost all social sciences rely on statistical methods to analyze data and to form hypotheses, and almost all of them use to a greater or lesser extent, a range of mathematical statistics methods to help us understand the social world. The increasing importance of mathematical statistics methods in the social sciences is the formation of new sub-disciplines, and the establishment of specialized journals and societies.

Corresponding author: **Dr. Deepak Kumar Routray,**
ISPAT Autonomous College, India.

[1] Social science" includes disciplines such as anthropology, political science, and sociology, but also economics and parts of linguistics and psychology.

There are similar debates in other social sciences, but it is important to stress that problem of one method such as axiomatization or the use of set theory which can hardly be taken as a sign of bankruptcy of mathematical statistics methods in the social sciences tout court. This paper surveys mathematical statistics methods used in the social sciences, significance test, the development of statistical model, need of statistical model in social sciences and discusses some of their philosophical questions.

Why need the Statistics in Social Sciences?: A historically important reason, why statistics is used in the social sciences was that `statistics' is associated with precision and objectivity. These are (arguably) two requirements any science should satisfy, and so the statistical mathematics in the social sciences was considered to be a crucial step that had to be taken to turn the social sciences into real science. Some such view has been defended by many authors. Luce and Suppes (1968), for example, argue along these lines for the importance of axiomatizing the theories of the social sciences. These authors also developed measurement theory (Krantz et al. 1971) and Suppes (1967, 2001) showed how the relation between a theory and its target system can be explicated in statistical terms. Contrary to this tradition, it has been argued that the subject matter of the social sciences does not require a high level of precision and that the social sciences are and should rather be inexact (cf. Hausman 1992).

After all, what works in the natural sciences may well not work in the social sciences. While Sir Karl Popper, one of the towering figures in the methodology of social science, did not promote the statistics in the social sciences in the first place (Hands 2008), it is clear that it nevertheless plays an enormous role in his philosophy. Given his focus on predictions and falsifiability, a theory that is statistzed is preferable to a theory that is not. After all, it is much easier to derive falsifiable conclusions from clearly stated propositions than from vague and informal claims. It is a mistake, however, to overestimate the role of the statistics. At the end, statistics provides the social scientist only with tools, and what one obtains when using these tools will crucially depend on the assumptions that are made. This is a variant of the well known GIGO principle from computer science (garbage in, garbage out). All assumptions are motivated informally; formulating them in the language of statistics just helps putting them more precisely. And once the assumptions are formulated statistically, the machinery of statistics helps to draw inferences in an automated way. This holds for analytical calculations as well as for numerical studies, including computer simulations (Frigg and Reiss 2010; Hartmann 1996).

This brings us to another advantage of mathematical statistics methods in the social sciences. While non-formal theories often remain rather simplistic and highly idealized, formal theories can be complicated and more realistic, rejecting the messiness of our world. The mathematical statistics machinery then helps drawing inferences which could not be obtained without them (Humphreys 2004). Often different assumptions pull in opposite directions, and it is not clear which one will be stronger in a specific situation. However, when implemented in a mathematical statistics model, it can be derived what happens in which part of the parameter space. And so the availability of powerful computers allows the systematic study of more realistic models. There is, however, also a danger associated with this apparent advantage. Given the availability of powerful computers, scientists may be tempted to construct very complex statistical models. But while these models may do well in terms of empirical adequacy, it is not so clear that they also provide understanding. This is often provided by rather simple models sometimes called `toy models`, i.e. models that pick only one crucial aspect of a system and help us to get a feel for what follows from it.

The Development of Statistical Model: Statistical methods are nowadays a central method of the social sciences. First, it is indispensable for evaluating experimental data e.g. in behavioural economics or experimental psychology. For instance, psychologists might want to find out whether men act, in a certain situation, differently from women, or whether there are causal relationships between violent video games and aggressive behaviour. Second, the social sciences heavily use statistical models as a modelling tool for analyzing empirical data and predicting future events, especially in econometrics and operational research, but recently, also in the mathematical branches of psychology, sociology, and the like. For example, time series and regression models relate a number of input (potential predictor) variables to output (predicted) variables. Sophisticated mathematical statistics model comparison procedures try to elicit the structure of the data-generating process, eliminate some variables from the model, select a best model" and finally fit the parameter values to the data. The conception of statistics as an inferential tool is quite young: throughout the 20th century, statistics was mainly used as a descriptive tool to summarize data and to fit various models. While in inferential statistics, the focus lies on testing hypotheses against each other, or quantifying evidence for or against a certain hypothesis, descriptive statistics focuses on summarizing data and fitting the parameters of a given model to a set of data. The most famous example is the method of the least squares, a procedure to centre a data set (Xn,Yn) where n € N around a straight line. Other important descriptive statistics are contingency tables, effect sizes, and measure of central tendency and dispersion measures. Descriptive statistics were, however, "statistics without probability" (Morgan 1987), or we can say statistics without uncertainty. In the late 19th and early 20th century, science was believed to be concerned with certainty, with the discovery of invariable, universal laws. This left no place for uncertain. We can recall that at that time, stochastic theories in the natural sciences, such as statistical mechanics, quantum physics, or laws of inheritance, were still quite new or not yet invented. Furthermore, there was a hope of reducing them to more fundamental, deterministic regularities, e.g. to take the stochastic nature of statistical mechanics as an expression of our imperfect knowledge, our uncertainty, and not as the fundamental regularities that govern the motion of molecules. Thus, statistical modelling contradicted the nomothetic ideal (Gigerenzer 1987), inspired by Newtonian and Laplace and physics, of establishing universal laws. Therefore statistics was considered as a mere auxiliary, imperfect device, a mere surrogate for proof by deduction or experiment. For instance, the famous analysis of variance (ANOVA) obtained its justification in the nomothetic view through its role in causal inference and elucidating causal laws.

It is interestingly says that these views were held even in the social sciences, although the latter dealt with a reality that was usually too complex to isolate causal factors in laboratory experiments. Controlling for external impacts and confounders poses special problems to the social sciences, whose domain are humans and not inanimate objects. The search for deterministic, universal laws in the social sciences might thus seem futile - and this is probably the received view today- but in the 1st half of the 21st century many social scientists thought differently. Statistics was needed to account for measurement errors and omitted causal influences in a model, but it was thought to play a merely provisional role:

> "Statistical devices are to be valued according to their efficacy in enabling us to lay bare the true relationship between the phenomena under consideration. An ideal method would eliminate all of the disturbing factors." (Schultz 1928, 33)

Thus, the view of statistics was eliminativist: as soon as it has done its job and elucidated the laws which we aim at, we can dismiss it. In other words, the research project consisted in eliminating probabilistic elements, instead of discovering statistical laws and regularities or modelling physical quantities as probabilistic variables with a certain distribution. This methodological presumption, taken from 19th century physics, continued to haunt social sciences far into the first half of the 20th century. Basing on the above, in total, there are three main reasons why inferential statistics was recognized as a central method of the social sciences:

- The advances in mathematical probability, as summarized by Kolmogorov in his work (1933/56).

- The inferential character of many scientific questions, e.g. whether there is a causal relationship between variables X and Y. There was a need for techniques of data analysis that ended up with an inference or a decision, rather than with a description of a correlation.
- The groundbreaking works by particular pioneer minds, such as Tinbergen and Haavelmo in economics (Morgan 1987).

Significance Tests and Statistical Decision Rules: One of the great conceptual inventions of the founding fathers of inferential statistics was the sampling distribution (Fisher 1935). In the traditional approach that is in classical regression, there was no need for the concept of a sample drawn from a larger population- instead, the modelling process directly linked the observed data to a probabilistic model. In the modern understanding, the actual data are just a sample drawn out of a much larger, hypothetical population about which we want to make an inference. The rationale for this view of data consists in the idea that scientific results need to be replicable. Therefore, we have to make an inference about the comprehensive population or the data-generating process for the matter, instead of making an `in-sample' inference whose validity is restricted to the particular data we observed. This idea of a sampling distribution proved crucial for what is known today as frequentist statistics. That approach strongly relies about this idea of the sampling distribution, outlined in the works of Fisher (1925, 1935, and 1956) and Neyman and Pearson (1933, 1967), parting ways with the classical accounts of Bayes, Laplace, Venn and others.

When we take the concept of frequentist statistics, we found there is a sharp division between approaches that focus on inductive behaviour, such as the Neyman-Pearson school, and those that focus on inductive inference, such as Fisherian statistics. To elucidate the difference, we will present both approaches in a nutshell. Neyman and Pearson (1933) developed a behavioral framework for deciding between two competing hypotheses. For instance, take the hypothesis H_0 that a certain learning device does not improve the student's performance, and compare it to the hypothesis H_1 that there is such an effect. The outcome of the test is interpreted as a judgment on the hypothesis, or the prescription to take a certain action ("accept/reject H_0"). They contrast two hypotheses H_0 and H_1 and develop testing procedures such that the probability of erroneously rejecting H_0 in favour of H_1 is bounded at a certain level "α", and that the probability of erroneously rejecting H_1 in favour of H_0 is given that constraint, as low as possible. In other words, Neyman and Pearson aim at maximizing power of a test that is the chance of a correct decision for H_1 under the condition that the level of the test that is the chance of an incorrect decision for H_1 is bounded at a real number "α". Thus, they developed a more or less symmetric framework for making a decision between competing hypotheses, with the aim of minimizing the chance of a wrong decision.

While such testing procedures apply well to issues of quality control in industrial manufacturing and the other, the famous biologist and statistician Ronald A. Fisher (1935, 1956) argued that they are not suitable for the use in science. First, a proper behavioristic, or decision-theoretic approach has to determine costs for faulty decisions which was done by Neyman-Pearson by choosing the level 'α' of a test. This involves, however, reference to the purposes to which we want to put our newly acquired knowledge. For Fisher, this is not compatible with the idea of science as pursuit of truth. Statistical inference has to be convincing to all freely reasoning minds, entirely independently of any intentions that might be furthered by utilizing the knowledge inferred" (Fisher 1956). Second, in science, a judgment on the truth of a hypothesis is usually not made on the basis of a single experiment. Instead, we obtain some provisional result which is refined through further analysis. By their behavioral rationale and by making a "decision" between two hypotheses, Neyman and Pearson insinuate that the actual data justify a judgment on whether H_0 or H_1 is true and accepted or rejected Such judgments have, according to Fisher, to be suspended until further experiments confirm the hypothesis, ideally using varying auxiliary assumptions and experimental designs. Third, Neyman and Pearson test a statistical hypothesis against a definite alternative. This leads to some results that appear paradoxical. Take, for instance, the example of a normal distribution with known variance $\sigma^2 = 1$ where the hypothesis about the mean $H_0 : \mu = 0$ is tested against the hypothesis $H_1 : \mu = 1$. If the average of the observations centres, say, around -5, it appears that neither H_0 nor H_1 should be `accepted'. Nevertheless, the Neyman-Pearson rationale contends that in such a situation we have to accept H_0 because the discrepancy to the actual data is less striking than with H_1. In such a situation, when H_0 offers a poor fit to the data, such a decision is arguably weird. Summing up, Fisher disqualifies Neyman and Pearson's decision-theoretic approach as a mathematical "reinterpretation" of his own significant tests that is utterly inappropriate for use in the social sciences ,he even suspects that Neyman and Pearson would not have come up with their approach had they had any real familiarity with work in the natural sciences" (Fisher 1956, 76). Therefore he developed a methodology of his own which proved to be extremely influential in the natural as well as in the social sciences. His books, "Statistical Methods for Research Workers" (1925) and "The Design of Experiments" (1935) shaped the applications of statistics in the social sciences for decades. The core of his method is the test of a point null hypothesis or significance test. Here, we want to tell chance effects from real effects. To this end, we check whether a null (default, chance) hypothesis is good enough to fit the data. For instance, we want to test the effects of a new learning device on students' performance, and we start with the default assumption that the new device yields no improvement. If that hypothesis is apparently incompatible with the data (if the results are `significant'), we conclude that there is some effect in the treatment. The core of the argument consists in `Fisher's Disjunction':

> "Either an exceptionally rare chance has occurred, or the theory [=the null hypothesis] is not true."(Fisher 1956, 39)

It is a mistake, however, to overestimate the role of the statistics. At the end statistics provides the social scientist only with tools, and what one obtains when using these tools will crucially depend on the assumptions that are made. This is a variant of the well known GIGO principle from computer science ("garbage in, garbage out). All assumptions are motivated informally; formulating them in the language of statistics just helps putting them more precisely. And once the assumptions are formulated mathematically, the machinery of statistics helps to draw inferences in an automated way. This holds for analytical calculations as well as for numerical studies, including computer simulations Frigg and Reiss (2010) Hartmann (1996).

This brings us to another advantage of mathematical statistics in the social sciences. While non-formal theories often remain rather simplistic and highly idealized, formal theories can be complicated and consider more realistic, rejecting the messiness of our social world. The statistical machinery then helps drawing inferences which could not be obtained without them (Humphreys 2004). Often different assumptions pull in opposite directions, and it is not clear which one will be stronger in a specific situation. However, when implemented in a mathematical statistics model, it can be derived what happens in which part of the parameter space. And so the availability of powerful computers allows the systematic study of more realistic statistical models.

Statistical versus practical significance: The null hypothesis typically denotes an idealized hypothesis, such as there is no difference between the effects of A and B". Practically no one believes such a hypothesis to be literally true, rather, everyone expects that there are differences, but perhaps just at a minute degree. The effects of A and B are always different in some decimal place for some A and B. Thus asking are the effects different?' is foolish." (Tukey 1991, 100) However, even experienced scientists often read tables in an article by looking out for asterisks: one asterisk denotes significant" findings ($p < 0{:}05$), two asterisks denote highly significant" ($p < 0{:}01$) findings. It is almost impossible to resist the psychological drive to forget about the subtle differences between statistical and scientific significance, and many writers exploit that fact. All psychologists know that statistically significant does not mean plain-English significant, but if one reads the literature, one often discovers that a finding reported in the Results sections studied with asterisks becomes in the discussion section highly significant or very highly significant, important, big!" (Cohen 1994, 1001) Instead, statistical significance should at best mean that evidence speaks against our idealized hypothesis, while we are still unable to give the direction of departure or the size of the observed effect (Kirk 1996). This provisional interpretation is in line with Fisher's own scepticism regarding the interpretation of significance tests, and Keuzenkamp and Magnus's (1995) observation that significance testing in econometrics rarely leads to the dismissal of an economic theory, and its subsequent replacement. Finally, under the assumption that null hypotheses are strictly spoken wrong, it is noteworthy that significance tests bound the probability of erroneously rejecting the null, while putting no constraints on the probability of erroneously accepting the null, i.e. the power of a test. Considerations of power, sample size and eject size that are fundamental in Neyman and Pearson's approach fall out of the simplified Fisherian picture of significance testing. This is not to say that these tests are worthless for instance, in econometrics, a series of significance tests can be very useful to detect whether a statistical model of a certain process has been misspecified. Significance tests look for directions in different departures (autocorrelation, moving average, etc.), and significant results provide us with reasons to believe that our model has been misspecified, and make us think harder about the right form of the model that we want to use in future research (Mayo and Spanos 2004; Spanos 1986). In that spirit, it should be stressed once more that Fisher considered significance tests to be a preliminary, explorative form of statistical analysis that gives rise to further investigations, not to final decisions on a hypothesis. But reading social science journals, it is not always clear that the practicing researchers are aware of the problem.

Concluding Remarks

Analysing the above facts and findings we concluded that, a variety of mathematical statistics methods that are used in the social sciences and argued that such techniques in spite of several methodological objections can add extra value to social scientific research. Then, we have focused on methodological issues in statistics the part of mathematical statistics that is most frequently used in the social sciences, in particular in the design and interpretation of experiments. We have represented the emergence of and the rationale behind the ubiquitous significance tests, as well as explained the pitfalls to which many researches fall prey when using them. Finally, after comparing significance testing to rivalling schools of statistical inference, we have discussed recent trends in the methodology of the social sciences and argued that there is reason for optimism, and that awareness of methodological problems, as well as interest for mathematical and statistical techniques is growing.

REFERENCES

Anderson, P., K. Arrow and D. Pines (eds.) 1988. The Economy As An Evolving Complex System. Redwood City: Addison-Wesley.

Armitage, P., and G. Berry 1987. Statistical Methods in Medical Research. Second Edition. New York: Springer.

Arrow, K. et al (eds.) (1960): Mathematical Methods in the Social Sciences.Stanford: Stanford University Press.

Backhouse, R. 1995. A History of Modern Economic Analysis. Oxford: Blackwell.

Balzer, W. and B. Hamminga (eds.) (1989): Philosophy of Economics. Dordrecht:Kluwer.

Balzer, W., C.U. Moulines and J. Sneed (1987): An Architectonic for Science: The Structuralist Program. Dordrecht: Reidel.

Beed, C. and O. Kane 1991. What Is the Critique of the Mathematization of Economics? Kyklos 44 (4), 581-612.

Beisbart, C. and S. Hartmann 2009. Welfarist Evaluations of Decision Rules under Interstate Utility Dependencies, Social Choice and Welfare doi: 10.1007/s00355-009-0399-z.

Berger, J.O., and T. Sellke (1987). Testing a point null hyphoteses: The irreconciliability of pvalues and evidence (with discussion). Journal of the American Statistical Association 82, 112-122.

Berm_udez, J.L. 2009. Decision Theory and Rationality. Oxford: Oxford University Press.

Casella, G., and R. L. Berger 1987. Reconciling Bayesian and Frequentist Evidence in the One-Sided Testing Problem. Journal of the American Statistical Association 82, 106-111.

Cohen, J. (1994): The Earth is round (p <:05). American Psychologist 49, 997-1001.

Debreu, G. 1991. The Mathematization of Economic Theory. American Economic Review 81(1),

Edwards, W., H. Lindman and L.J. Savage (1963): Bayesian Statistical Inference for Psychological Research. Psychological Review, 70, 450-499.

Fisher, R.A. 1925. Statistical Methods for Research Workers. Edinburgh: Oliver and Boyd.

Fisher, R.A. 1935. The Design of Experiments. Edinburgh: Oliver and Boyd.

Fisher, R.A. 1956. Statistical Methods and Scienti_c Inference. New York: Hafner.

Fogel, R.W. 1975. The Limits of Quantitative Methods in History, American Historical Review 80(2): 329-50.

Frigg, R. and J. Reiss 2009. The Philosophy of Simulation: Hot New Issues or Same Old Stew? Synthese 169 (3).

Frigg, R. and S. Hartmann: Models in Science, in: The Stanford Encyclopedia of Philosophy, (Spring 2006 Edition).

Gaertner, W. (2006): A Primer in Social Choice Theory. New York: Oxford University Press.

Gallegatti, M., S. Keen, T. Lux, and P. Ormerod 2006. Worrying Trends in Econophysics, Physica A. Physica A 370(1), 1-6.

Gigerenzer, G. 1987. Probabilistic Thinking and the Fight against Subjectivity, in Kruger et al. (eds., 1987): The Probabilistic Revolution, 11-33.

Gigerenzer, G. 1993. The Superego, the Ego, and the Id in statistical reasoning. In G. Keren & C. Lewis (Eds.), A handbook for data analysis in the behavioral sciences: Methodological issues, 311-339. Hillsdale, NJ: Erlbaum.

Gigerenzer, G. 2008. Rationality for Mortals: How Humans Cope with Uncertainty. Oxford: Oxford University Press.

Gilbert, N. and K. Troitzsch 2005. Simulation for the Social Scientist. New York: McGraw-Hill.

Goodman, S. 1999. Toward Evidence-Based Medical Statistics. 1: The P Value Fallacy. Annals of Internal Medicine 130,995-1004.

Granger C. 1999. Empirical Modelling in Economics: Speci_cation and Evaluation. Cambridge: Cambridge University Press.

Grodon, S. 1991. The History and Philosophy of Social Science. London: Routledge.

Harlow, L.L., S.A. Mulaik, and J.H. Steiger (eds.) 1997. What if there were no significance tests? Mahwah/NJ: Erlbaum.

www.google.com.

****