PEER REVIEW JOURNAL
FOR
RESEARCH AND READINGS IN
APPLIED STATISTICS

ISSN 2321-0877

SANKHYA VIGNAN
સંખ્યા વિજ્ઞાન

NEW SERIES (NSV 18) JUNE 2022 NO. 2

Published by
GUJARAT STATISTICAL ASSOCIATION

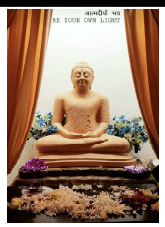
http://www.sankhyavignan.org

SANKHYA VIGNAN
સંખ્યા વિજ્ઞાન

(e-mail ID : svgsa2015@gmail.com)

EDITORIAL BOARD

EDITORIAL



We are extremely happy to bring before you this issue (NSV 18, June 2022, No. 2). We express our sincere thanks to all our contributors, evaluators, readers and well wishers for their continuous and consistent support, without which we would not have achieved our goal.

This issue contains for the section of Management and statistics, four articles. Other portion of the journal contains three research articles, one review article, one Biography, two book reviews and other sections like SV News letter and Readers form as usual. You will find the details int he content.

We are very highly obliged to our following referees who have helped us for evaluation of articles / papers submitted for this issue.

(Their names are given one by one in the order of their appearance in the journal.)

(1) D. S. Dave	**(2) J. R. Purohit**
(3) M. B. Thaker	**(4) V. H. Bajaj**
(5) P. P. Prajapati	**(6) C. D. Bhavsar**
(7) H. D. Budhbhatti	**(8) R. G. Bhatt**
(9) Shailesh Teredesai	

You can visit our website **www.sankhyavignan.org** and give your feedback and suggestions.

We express our sincere thanks to our **Research Team** for this work. In particular we thank **Shree Dinesh Darji for DTP work and Shree Ashish Bhatt for website.**

All our contributors will get digital copy and official certificates.

We express our warm wishes for your good health, progress and prosperity.

HAPPY STATISTICS DAY.

AHMEDABAD — **EDITORIAL BOARD**

Date : 29/06/2022

NOTE : Members of editorial board are in no way concerened with the views, opinions or ideas expressed in this issue. Authenticity responsibility lies solely with the persons presenting them.

SANKHYA VIGNAN

PEER REVIEWED REFREED QUARTERLY JOURNAL (ISSN : 2321-0877)

(Journal of Research and Readings in Applied Statistics)

*** Listed at International ISSN Directory, PARIS**

Average Circulation Rate : 850

Average Journal Evaluation Score : 7.50 (As on 31st March, 2022)

PUBLISHED BY GSA

GUJARAT STATISTICAL ASSOCIATION - FROM EDITOR'S DESK

SANKHYA VIGNAN is a peer reviewed refereed Bi-Annually journal that published empirical, conceptual and review papers of exceptional quality that contribute to Statistics Theory and enriched Applications of Statistical Techniques in various fields. The objective ofthe Journal is to disseminate knowledge, which ensures good practice of professional management and its focal point is on research and reflections relevant to academicians and practitioners in the field of Applied Statistics.

PEER REVIEW PROCESS

(1) To maintain PEER REVIEW REGISTER giving full details about authors, Refrees, decisions and internal rankings, before printing the Issue.

(2) All research articles published in the SANKHYA VIGNAN will undergo full peer review process. It should be clear for authors that the Editor In Chief is responsible for the final decision about the submitted papers; have the right to accept/reject any paper. The editor In Chief will choose any option fromt he following to reveiw the submitted papers:
   a) send the paper to two reviews, if the results were negative by one reviewer and positive by the other one; then the editor may send the paper for third reviewer or he take immediately the final decision by accepting/rejecting the paper. The Editor In Chielf will ask the selected reviewers to present the results within 7 working days, if they were unable to complete the review within the agreed period then the editor have the right to resend the papers for new reviewers using the same procedure. If the Editor In Chief was not able to find suitable reviewers for certain papers then he has the right to accept/reject the paper.
   b) send the paper to a selected editorial board member(s) or reviewer.
   c) the Editor In Chief himself evaluates the paper.

(3) The authors of the papers/articles accepted for publication need to submit a duly signed copyright form within maximum one week of their receiving intimation of acceptance of their paper/article.

(4) To give each published issue for evaluation by two external Refrees other than editors, obtain evaluation reports and journal evaluation score out of 10 points.

(5) To declare average Journal Evaluation Scaore on the basis of the scores given by these two Refrees for each published issue.

(6) To keep all records in perfect order.

(7) Invite and accept suggestions made by the Refrees for Peer Review.

PEER REVIEW COMMITEE:

It consists of six members :

*** Editor In Chief** *** Two Joint Editors** *** Two External Referees** *** Editorial Secretary**

PAPERS/MANUSCRIPT EVALUATION CRITERIA

Each submitted manuscript is evaluated on the following basis :

- The originality of its contribution.
- The soundness of its theoryu and methodology.
- The coherence of its analysis.
- The ability to communicate to readers (grammar and style)
- All research paper are reviewed by at least two suitable qualified experts.
- All publication decisions are made by the journals' Editor on the basis of the review provided.
- Normal turn-around time for evaluation of manuscripts is 21 days from the date of recipient.

GUIDELINES FOR AUTHORS

1. The first page of the article should contain full name of authors with designation and correspondence address, contact numbers and email ID's (Both Institution & residential)
2. The research must be original in nature and is neither published nor under consideration for publication anywhere else.
3. The review process may take 2-3 days and the status would be known within 24 hours of submissions of the article. It is mandatory to mention a valid email and mobile number. We acknowledge the receipt of your article and subscription fee be email.
4. The abstract s hould not be more than 200-2502 words and it should adequately describe the work and highlight its signifcance. Avoid the use of abbreviation and reference in the abstract. The abstract should be folowed by relevant keywords.
5. The main paper (Manuscripts) should not exceed more than 3,000 to 4,000 words (inclduing graphs & charts.)
6. Article should be typed in 12 point - Times New Roman Font English with a one and half space and single column with 1 Margin on a Standard A4 size paper. The page numbers should be at the center of every page. All headings & sub headings must be in bold letters.
7. Table should be numbered consecutively, the title of the table should be placed above the table. The source should be indicated at the bottom.
8. All the tables, charts, graphs, diagrams should be in black and not in colors.
9. Footnotes, italics and quotation marks should be kept to the minimum.
10. References should be mentioned in APA Referencing Formate.

HOW TO SUBMIT

A) We will accept soft copies of article through online submissions at the Email ID: (1) svgsa2045@gmail.com, (ii) drjayesh.purohit@gmail.com, (iii) bbjani2012@gmail.com

B) Two hard copies should be sen to Dr. B. B. Jani, Chief Editor, Sankhya Vignan, B/14, Bansidhar Apartment, Mirambica Road, Naranpura, Ahmedabad - 380 013. Ph. : (R) 079-2747 6770 (M) 98240 57902

MANAGEMENT AND STATISTICS
(ARTICLE)

RESEARCH MISCONDUCT AND RESEARCH FRAUD

A. C. Brahmbhatt*

ABSTRACT

This paper describes briefly about the academic and cultural norms pertaining to all types of research works that are carried out. There is also a caution as observed by the author to regularise the quality maintenance for the academic materials submitted for research.

KEYWORDS

Fabrication, Falsification, Plagiarism, Data Fishing

A plethora of research work is being carried out in social sciences, humanities, management, physical sciences , medical sciences , pharmacology, microbiology and virology , space science etc. world over.Vast majority of the researchers engaged in these fields is fundamentally honest and although the corners may be cut , the intention to add to the existing body of knowledge is beyond any doubt , nor would there be any effort to deceive the vast readership or the fund providers for research. Still in reality, we experience that there are researchers who indulge in misconduct and research fraud and bring blemish to the fair name of research. Green Steel (2011) analysed article retractions and observed that the research fraud has considerably increased in the recent years.

Research misconduct and fraud have spread on a wide spectrum and manifest in dishonest behaviors ranging from cherry-picking of data to approving of a chosen hypothesis to outright fabrication. The Federal Office of Research Integrity (2017) defines mainly three types of research misconduct: fabrication, falsification and

* Research Mentor, PDPU, Gandhinagar, Email : acbpramukh@hotmail.com
Author is thankful to the referee for reviewing.
(rcd. June '22 / rvd. June '22)

plagiarism.

FABRICATION : it means trying very hard to produce the ' desirable' results through data analysis .Fabrication is not a new phenomenon, it has a long history, in fact it existed since the dawn of sciences (Hensen & Miller, 1992). The most extreme form is that the research has been entirely fabricated. A Harvard University psychologist fabricated and falsified data and made false statements about experimental methods in 6 federally funded studies.It was ultimately unearthed by the Office of Research Integrity of the US department of Health and Human Sciences, in September , 2012. He had resigned for research misconduct in 2011.

FALSIFICATION: It refers to manipulating research materials , equipment or process or omitting data or results such that the research is not accurately represented in the research record. It can happen prior to, during or after data collection. In practice it is observed that much of it happens during data processing, statistical analysis and representation of the results.

PLAGIARISM: If the publication includes substantial sections from other researchers' work without any citation or acknowledgement of the original source, it is plagiarism.Some people are simply in the habit of stealing from other peoples' work and showing in their names on a continual basis. They are called the serial plagiarists. Out of 26 papers published by a professor of literature, at the University of Nevada, Las Vegas, 23 were found to contain several plagiarized passages ! The highest ethical standard does not approve even the self plagiarism.

In addition to this, there are various other types of research misconduct and unfair practices.

NOT TO DO RESEARCH AT ALL: There are so called researchers who first manage to take grant from a funding agency and then do not conduct the research at all, on which their published articles are supposedly based.One professor of political anthropology from the Vrije UNiversteit , Netherland had never published 61 of the papers listed on his CV.

ALTERED DATA: Some researchers do not indulge in complete falsification

, but make convenient adjustments in the data to arrive at the desired conclusions. They do not hesitate in altering the facts about subjects to , any how reach the desired findings. In one case, a researcher established the correlation between autism and measles-mumps and rubella vaccine. People trusting the research , refrained from administering it to their children. It resulted in the widespread outbreaks of measles and mumps in Europe and US.

DUPLICATION OF PUBLICATION: Some researchers submit the same articles/papers to more than one journal simultaneously . It is also a part of research misconduct.

SALAMI SLICING OF RESEARCH PAPER: Many researchers publish several articles from the same piece of research. It is not so easily detected and gives credit to authors misleading the academic community. Taking the same data twice into result calculations can significantly distort the final outcome of meta analysis.(Smoicic, 2013).

DATA FISHING : As a fisherman uses multiple hooks to catch the big fish, some researchers go on employing various statistical tools and techniques until they get statistically significant result.

DATA TRIMMING: It is another unfair practice in analyzing research data that happens when the researcher selectively abandons part of the data that is considered invalid , as it may not have been collected according to the protocol. For example removing data contributed by unqualified respondents after data collection is complete. It introduces substantive biases in the statistical analysis as such actions fail to ensure the randomness of the data set.

REPRESENTING AND INTERPRETING DATA: Same data can be represented in drastically different ways , which in turn can be used to support contrary interpretations. In his classic book ,' How to lie with Statistics ' Huff (1954), shares numerous examples of manipulating statistical analysis to support conclusions that are different from or even opposite to what the data really dictates. There are number

of mathematical gimmicks that are used to distort interpretation , e.g. unrepresentative samples, selective use of mean or median as 'average' , selective use of percentages, inadequate sample size, ignoring the assumptions and limitations of different statistical tests and techniques.

Conflict of Interest: Researchers have to play multiple roles at the same time. These roles include those of university employees, Ph.D. supervisors, teachers, mentors, consultants to governmental agencies or corporate units etc. Each of the roles might bring with it particular obligations to different groups and at times these obligations might compete or conflict with each other. For example, university professor who is also consultant to a company might be asked by the company to delay the publication of research so that it can have a competitive advantage. Delaying publication might negatively impact the image of his/her university. Delaying research results also frustrate the entire academic community as it anxiously awaits to learn about most recent research findings.

The most common causes for researchers getting indulged in such unfair practices are –i) temptation to get research funding ii) institutional pressure for' publish or perish '. iii) fad for publishing positive results.

REFERENCES:

[1] **Green Steel (2011),** Retractions in the scientific literature: Is the incidence of research fraud increasing? , Journal of Medical Ethics , 37(4), pp.249-53

[2] **Hersen , M, & Miller D.J. (Eds,.) (1992) ,** Future Directions; a modest proposal, Research fraud in the Behavioral and Biometrical Sciences, pp. 225-43 (New York, Wiley).

[3] **Federal Office of Research Integrity (2017),** Definition of Research Misconduct, https:// ori.hss.gov/definition-misconduct.

[4] **Smiocic V.S. (2013),** Salami Publication; definition and examples. Biochemia medica, 23 (3), pp.37-41.

[5] **Huff, Darrel (1954),** How yo lie with Statistics ? Norton W.W. & Co. Inc.

MANAGEMENT AND STATISTICS
(ARTICLE)

PARADIGM SHIFT IN BUILDING SUPPLY CHAINS : IMPACT OF ON GOING RUSSIA / UKRAINE CONFLICT

Ajay K. Aggarwal[(1)], Varinder M. Sharma[(2)] and Dinesh S. Dave[(3)]

ABSTRACT

This paper discusses briefly about the supply chain scenerio with its global impact arising due to the recent war situation between Ukrain and Russia. It provides a caution to the world about the seviour situation arising due to the present war situation.

KEY WORDS

PARADIGM SHIFT, GLOBAL SUPPLY CHAIN, CYBER SECURITY

OBSERVATIONS AND DISCUSSION

World has become interlinked and interdependent has become a cliche´. More than sharing borders, culture, and information, the western nations have become increasingly dependent upon developing nations as production locations to meet their product needs through Global Supply Chains (GSC) that have become a convenient, cost-efficient, and reliable *modus operandi* to serve the west (Simchi-Levi and Haren, 2022). The resilience of GSC has generally depended upon their ability to withstand external challenges like global warming, economic swings, political changes, tech-

(1) Professor of Management School of Business, Henderson State University, Arkadelphia, Arkansas 71999 (USA)
(2) Professor of Marketing, Dept. of Marketing, Eberly College of Business, Indiana University of Pennsylvania, Indiana, Pennsylvania 15705 (USA)
(3) Director and Professor of SCM, Dept. of Marketing and SCM, John A. Walker College of Business Applachian State University Boone, North Calolina 28608 (USA)
email : daveds@appstate.edu
(rcd. March '22 / rvd. June '22)

nological challenges, and cultural and language barriers, among others (Liuima, 2022 and Swanson, 2022). However, a few successive recent events, COVID-19 and Russia/ Ukraine conflict are compelling scholars and managers to take a fresh look at the resilience of extant GSPs

The covid-19 pandemic created supply chain failures around the globe during 2019-21. The disruptions were widespread, from sourcing countries to distribution centers to retail stores. The pandemic exposed the drawbacks of lean GSC (see Roelofste, 2022 and Simchi-Levi and Haren, 2022). The efficient, optimized, supply chain models simply failed to deliver. Specifically, insufficient inventories to sustain prolonged supply shocks led to empty store shelves around the world (Segal, 2022 and Simchi-Levi and Haren, 2022). The world recognized the drawbacks of overdependence on dominant suppliers since 2008 financial crisis (see Simchi-Levi and Haren, 2022), however, it experienced limited success in finding alternative regional or local sourcing options with necessary technology, materials, and workers. The Russia/Ukraine conflict has essentially witnessed similar failures of traditional GSP models; several industries, such as energy, oil, precious metals, and agricultural are increasingly facing unreliable product delivery and inflated costs.

The mantra of minimal inventory, with limited safety stock across various supply chain stages produced cost effective delivery mechanism in a relatively predictable, stable world (see Simchi-Levi and Haren, 2022; Segal, 2022; and Swanson, 2022) Even with minimal or localized disruptions or short-lived, the resilience of the GSP models was mostly assured. For instance, natural disasters such as floods or forest fires, while disruptive, were often short lived, and mostly written-off as once in a lifetime event. GSP Risks were generally manageable. With low transportation costs and product production costs essentially dictated the selection of sourcing country, regardless of its location or political stature. The lingering effects of COVID-19 pandemic and those of recent eruption of Russia/Ukraine conflict have once again questioned the efficacy of GSPs in delivering the products necessary for daily

consumption causing shortages, inflated prices or at worse, non-availability in many parts of the world including the Western countries. With the continuous arrival of new variants of COVID-19 and emergence of hot spots of future conflicts, the macroenvironment of extant GSPs has become turbulent and such disruptions are expected to grow in the future (see Simchi-Levi and Haren, 2022; Roelofsten, 2022; Segal, 2022; and Liuima, 2022). The GSP models need reconfiguring to cope with this new reality. In fact, companies across the globe are crafting plans to deal with the current and evolving issues, and making several necessary, and some immediate changes in supply chain management practices (see Kilpatrick, 2022).

Since the start of Russia/Ukraine conflict, virtually all GSC links in and around Russia and Ukraine have ceased to operate, disrupting supplies, delaying, reducing, or canceling production around the world. To make matters worse, Russia has been subjected to crippling sanctions – the worst since the cold war by the western countries including the US, Europe and Asian countries like Australia and Japan. The global impact has been highest on energy, agriculture, and transportation industries; other affected industries include metals, automotive, and high-tech.

The United Nations data indicates that prior to the Russia/Ukraine conflict, a quarter of the world's wheat, one-third of the barley, and about two-thirds of sunflower-oil were sourced out of Russia and Ukraine. In addition to providing one-eighth of the world's oil and about one-fifths of its natural gas, Russia was a supplier of critical minerals (e.g. platinum-group elements), neon (used in semiconductor chip production), lead fertilizer supplier, and provider of railway infrastructure for Europe and Asia (see Molly, 2022).

As a result of Russian isolation from world-trade by the imposed sanctions is having global consequences not only on the availability and prices of petroleum-related products, food, metal, but has also escalated inflationary pressures on almost everything dependent on them (e.g., batteries, all shipped goods, all vehicles, etc.). Since few supplies, if any, can leave Ukraine due to incapacitating of Ukraine ports

for trade, the lack of essential wiring and electronic components for Europe's automotive industry has stymied production. Disruption of the Russian and Ukrainian sea routes and air routes has also impacted neighboring countries such as Romania and Bulgaria's sea ports air cargo markets. As a result of these disruptions, two trends appear to be taking shape. First, more companies are increasingly looking favorably upon producing goods domestically and taking step toward increased self-reliance. When the inability to produce domestically propels firms to pursue international suppliers, the preferred destinations are friendly countries, and collaborations there are typically sought with multiple suppliers. Second, there is a renewed urgency for securing additional critical inventory and capacity. At the same time, companies are taking special care in choosing alternate sourcing locations for key commodities. Artificial intelligence and advanced analytics are increasingly being employed to help provide better data visibility, suggest prescriptive measures for alerts, and help keep risks under control. High degree of unrest among consumers caused by persistent inflation has reignited the need for supply chain managers to stay cognizant of pricing issues. In particular, strategies need to be formed to deal with rising costs while maintaining profitability. There is also an increasing need to invest in cyber security, guard digital assets, and embrace supply chain risk from a holistic perspective.

CONCLUDING REMARKS

The lingering effect of Russia/Ukraine conflict has exacerbated challenges for the post-pandemic GSC efficacy and risk. Companies around the world are reconfiguring their supply chains to deal with shortages of supplies emerging from the conflict affected locations. GSC scholars should develop appropriate models to facilitate companies in designing resilient supply chains to cope with current and future external threats.

ACKNOWLEDGEMENT

Authors thank the referee for reviewing our paper.

REFERENCES

[1] **Callahan, Molly,** "The Russian-Ukraine War could throw global supply chains into chaos. And that is the best-case scenario," News at Northeastern University, Boston, MA, March 18, 2022, https://news.northeastern.edu.

[2] **Kilpatrick, Jim,** "Supply Chain Implications of the Russia-Ukraine Conflict," Deloitte Insight, March 25, 2022, https://www2.deloitte.com.

[3] **Liuima, Justinas,** "Key Supply Chain Risks from Russia's Invasion of Ukraine," Euromonitor International, March 17, 2022, https://www.euromonitor.com.

[4] **Roelofsten, Volker,** "How to Mitigate the Impact of the Russia-Ukraine War on Your Global Supply Chain," Supply Chain 247, April 9, 2022, https://www.supplychain247.com.

[5] **Segal, Edward,** "Supply Chain Crisis Worsens as Russia's War Against Ukraine Continues", Forbes, April 2, 2022, https://www.forbes.com.

[6] **Simchi-Levi, David & Haren, Pierre,** "How the War in Ukraine is Further Disrupting Global Supply Chains," Harvard Business Review, March 17, 2022, https://hbr.org.

[7] **Swanson, Ana,** "Ukrainian Invasion Adds to Chaos for Global Supply Chains," New York Times, March 1, 2022, https://www.nytimes.com.

HEALTH CARE ANALYTICS

Parag Shah*

ABSTRACT

Healthcare 2.0 will shift the healthcare ecosystem towards innovative care delivery models because of Healthcare Analytics which comprises of deployment of Artificial Intelligence, Internet of Things and advanced Data Analytics and more. The main aim of this paper is to provide a deep analysis on the research field of healthcare data analytics. The paper has listed some data analytics tools and techniques that have been used to improve healthcare performance in many areas such as: medical operations, reports, decision making, and prediction and prevention system.

KEYWORDS

Healthcare Analytics, Descriptive Analytics, Diagnostic Analytics, Predictive Analytics, Prescriptive Analytics

INTRODUCTION

Healthcare has become one of India's largest sectors, both in terms of revenue and employment. Healthcare comprises hospitals, medical devices, clinical trials, outsourcing, telemedicine, medical tourism, health insurance and medical equipment. The Indian healthcare sector is growing at a brisk pace due to its strengthening coverage, services and increasing expenditure by public as well private players. The

* Dept. of Statistics, H. L. College of Commerce,
Gujarat University, Ahmedabad-380 009. email : pbshah@hlcollege.edu
(rcd. March '22 / rvd. June '22)

Government of India is planning to increase public health spending to 2.5 per cent of the country's GDP by 2025. India has also become one of the leading hubs for high-end diagnostic services with tremendous capital investment for advanced diagnostic and medical facilities at a very low cost compared to other countries.

Healthcare 1.0 was broadly defined by a focus on defensive medicine, billing, and fee-for-service, culminating in the mass adoption of EMRs and data proliferation. Healthcare 2.0 is a new wave focused on improving clinical efficiency, quality of care, affordability, and fee-for-value – culminating in a new age of healthcare analytics.

The healthcare analytics market in India is expected to reach a value of INR 47.04 Bn by 2025, expanding at a CAGR of ~20.49% during the 2020-2025 period. Factors like an increased focus on collection and analysis of data from different healthcare sources for improved customer service, technological advancements and rising adoption of electronic health records are expected to drive the growth of the healthcare analytics market in India. Analytics is rapidly gaining power as it helps healthcare systems reduce healthcare costs, improve quality of care and facilitate preventive care.

WHAT EXACTLY IS HEALTH CARE ANALYTICS?

Healthcare analytics is the branch of analysis that emphases on offering insights into hospital management, patient records, costs, diagnoses, and more. The field covers a broad band of the healthcare industry, offering insights on both the macro and micro level.

Healthcare analytics also integrates business intelligence suites and data visualization tools to reveal and understand historical data patterns, predict future events, and provide actionable insights to make fact-based decisions and improve clinical, financial and operational performance of healthcare organizations.

DATA ANALYTICS IN HEALTH CARE

This are some of the skills and techniques that will allow us to connect the potential of data analytics in Health Care.

- Correlation & Regression analysis
- Hypothesis testing
- Statistical Process Control
- Probability distributions
- Factor Analysis
- Cluster Analysis
- Patient Profiling
- Patient Segmentation
- Structural Equation Modeling
- Machine Learning
- Neural Network analysis
- Natural Language Processing
- Image recognition and
- Speech analysis
- Data mining
- Artificial Intelligence

WHAT TO EXPECT FROM HEALTHCARE ANALYTICS?

HealthCare analytics can be broadly classified into descriptive, diagnostic, predictive, and prescriptive analytics.

- **Descriptive analytics: What has happened?**

  In its most basic form, healthcare data only tells you what has already happened. How many patients were admitted to the hospital last year? How many were discharged within a week? Average billing amount per patient?

  Descriptive analytics is the ability to quantify events and report on them in a

human-readable way using tables and charts. It's the first step in turning big data into actionable insights, and there is a lot to be learned from this level of analytics.

- **Diagnostic analytics: Why it happened?**

  It takes descriptive data a step further and provides deeper analysis to answer the question: Why did this happen? Often, diagnostic analysis is referred to as root cause analysis. This includes using methods such as data discovery, data mining, and drill down and drill through.

  Diagnostic data analytics works backward from the symptoms, explore data, make correlation & association to suggest the cause of what has happened. While doctors continue to be responsible for the final diagnosis, they can use this data analytics to save time and to avoid possible errors of judgment.

- **Predictive analytics: What might happen in the future?**

  It uses information from diagnostic and descriptive analytics to predict the future events. For example, which patients will have the highest risk of hospitalization next week? Which person is more prone to heart disease based on his medical history? Which patient may default in paying hospital bills?

  In predictive analytics, statisticians & data scientists use historical data to train models to predict future events by employing advanced statistical and computational techniques.

- **Prescriptive analytics: What needs to be done?**

  Prescriptive analytics is the most advanced healthcare analytic because it allows us to make specific recommendations on patient care delivery interventions based on the predictive model. It recommends specific actions in response to individual patient symptoms and medical history. The benefits of prescriptive analytics range from the identifying areas of improvement in treatment and protocols to reducing the rate of re-admitted patients, and lowering the cost of healthcare in general-from patient bills to the cost of operations in hospital billing departments, government expenditure.

CHALLENGES TO HEALTHCARE ANALYTICS IN INDIA

Though there are tremendous growth opportunity in this field, there are a set of challenges in implementation of HealthCare Analytics:

- Skilled workforce
- Lack of Digitization
- Complex and Incomplete data
- Sharing of data
- High investment Cost
- Establishing standards and governance
- Data Privacy & Security
- Overdependence on Public healthcare system

CONCLUSION

HealthCare Analytics encompasses mainly three types of analytics solutions, namely descriptive analytics, predictive analytics and prescriptive analytics. Healthcare practitioners, healthcare providers, patients, the government and pharmaceutical companies are the major end-users of healthcare analytics solutions. As in any nascent industry, adoption will be gradual, beginning with early adopters and then to mass market. But there is no doubt that the field of health care analytics is one whose time has come and will create immense value to the entire healthcare ecosystem in the next decade.

Companies like IBM Corporation, McKesson Corporation, UnitedHealthcare Group, Oracle Corporation, Allscripts Healthcare Solutions, Wipro etc. provide professional healthcare analytics services. But still there is a dire need for more such start-ups and companies in India.

The talent needed in country to leverage data analytics is in limited supply. Medical & pharmacy researchers, Statisticians, data scientist, biostatisticians could begin to use data and machine learning to produce game-changing insights and unlock

the value of large structured and unstructured data. Healthcare Analytics is sure to become a vibrant new career opportunity in India

ACKNOLWEDGEMENT

Authors thank the referee for reviewing this paper.

REFERENCES

[1] **Al Khatib et al (2015)** Analysis of Research in Healthcare Data Analytics, Australasian Conference on Information Systems

[2] **Lidong Wang & Cheryl Ann Alexander (2019),** Big Data Analytics in Healthcare Systems, International Journal of Mathematical, Engineering and Management Sciences, Vol. 4, No. 1, 17–26

[3] **Maria Mohammad Yusuf (2021)** Big Data Analytics in Health Care – A Review Paper, International Journal of Computer Science ad Information Technology, Vol. 2

[4] **Kornelia Batko & Andrzej Œlêzak (2022)** The use of Big Data Analytics in healthcare, Journal of Big Data, Springer 9(3)

[5] **Sabyasachi Dash et. al. (2019)** Big data in healthcare: management, analysis and future prospects, Journal of Big Data, Springer 6(54).

[6] **The Role of Data Analytics in Health Care (2021),** School of Health and Rehabilitation Sciences, University of Pittsburgh

MANAGEMENT AND STATISTICS
(ARTICLE)

REPLICATION CRISIS : A PHILOSOPHICAL STANDPOINT

Bhavya Patel*

ABSTRACT

The "replication crisis" stems from the inability of published scientific findings to hold confirmatory power whilst constantly undermining the credibility of scientific research. This paper examines this crisis from a philosophical standpoint.

KEYWORDS

Epistemology, p-hacking, HARKing

1. INTRODUCTION:

Replication is essential in science to validate research as an operationalization of objectivity and demonstrate that the scientific knowledge obtained from a particular experiment can be separated from the circumstances in which the experiment was conducted. Failure of replication casts doubt on an entire subfield and introduces the possibility that former findings decrease in robustness overtime (Brandt, 2014). An example of the functionality of replication includes a follow-up of the classic textbook experiment, the marshmallow experiment, in the late '80s which showed that children who could avoid eating the marshmallow, had higher SAT scores and fewer behavioral problems in their lifetime. However, it was only due to replication by Tyler Watts in the '90s, that it became evident that the participants' response is dependent on their intelligence, family background and, epigenetic alterations such that there is no

* Graduate Student, Neuroscience and Physiology (H.B.Sc.) University of Toronto, Toronto, Canda. Email : patelbhavyaa@gmail.com
(rcd. / rvd.)

correlation between delaying gratification and behavioral outcomes later in life. Popper, highlighting the necessity for replicability, states that "non-replicable single occurrences are of no significance to science (1959)."/

2. RATIONALE

The replication crisis has many origins, the foremost being publication bias, in which the outcome of a study may be manipulated to meet the criteria for publication. As the primary criterion for publication is statistical significance, questionable research practices such as p-hacking occur to control the flexibility of data collection to obtain favorable results for publication. Simmons et al., 2011 have demonstrated that p-hacking alone can increase the rate of false positives to 61%. Biased literature results from publishing only significant findings; 5% of studies are false positives in comparison to 95% of the true negatives that do not get published and remain in filing. Similarly, HARKing, a concept coined by Norbert Kerr, refers to a QRP that involves hypothesizing the outcome of a study after the results have been known which further complicates the ongoing crisis. Along with the proliferation of null hypothesis significance testing, the reward system of science contributes to the replication dilemma such that only the first scientist is rewarded in discovery, and this may pressurize some scientists to participate in scientific research that is non-replicable. The replication crisis raises the need to efficiently validate scientific research, and it becomes the effort of philosophers to provide a rational solution for the scientific community. This essay attempts to present the philosophical concerns regarding the replication crisis including the direct and conceptual dichotomy, the self-corrective thesis, and the variety of evidence thesis./

The greatest concern for philosophers is to/*define*/"replication." Machery, 2019 suggests, "Experiment A replicates Experiment B if and only if A consists of a sequence of events of the same type as B while resampling some of its experimental components to assess the reliability of the original experiment."/The recipe for

confirming empiric data raises two questions; which conditions need to be fulfilled to ensure replicability of Experiment B and what factors make it reliable that Experiment A is a successful replication of Experiment B. However, these questions are unanswerable, as ironically, there is a lack of research to address the validity of performing a replication experiment. Epistemological inquiries are raised by philosophers concerning the epistemic function of replication. This debate questions whether it would be accurate to claim that scientific discovery is reliable because it has been replicated than one that cannot be? Would it be valid to claim that non-replicable research has no epistemic value? However, it becomes evident that no study can be fully replicated and hence, furthers the inquiry of what constitutes a successful replication versus an unsuccessful one. Furthermore, if a study fails to replicate, that depends heavily on the definition and factors contributing to what replication success exactly constitutes and hence, becomes controversial (Romero, 2019).

To amend this disadvantage, philosophers use the term/"*direct replication*"/for experiments that are replicated by the creation of an identical experimental design to the original and/"*conceptual replication*"/that works to evaluate the generalizability of an experiment via modification of the original experimental design (Schmidt, 2009). As mentioned above, no experiment can be fully replicated, and this contradicts the idea of replication itself such that Experiment B is replicated by a different Experiment A (Schmidt, 2009). As such, both these notions are vague, and each has their opportunities and obstacles depending on the experiment that is being replicated.

Philosophers have made many attempts to re-define this dichotomy. An example includes viewing the differences between an original experiment and a replicated experiment via a matter of degree. Brandt et al, 2014 has attempted to create a "36 question replication recipe" for par excellence completion of a/*"close replication"*/ with ingredients ranging from high statistical power, defining the effects and methods of manipulation, and critically comparing replication results to the original study whilst making replication attempts as *close* as possible to the original experiment and

accepting unavoidable differences. Similarly, LeBel et al., 2019 identifies a scale of five types of replications based on methodological similarity to the original experiment./

Next, philosophers seek to understand the superior alternative between direct and conceptual replications. Some philosophers believe that conducting replications serves an epistemic goal, but some authors think otherwise. As minor changes in setting, timing, weather, and factors relating to participants such as ethnicity, height, and weight, can never be manipulated to be identical to the previous trial, some philosophers argue that a direct replication is not sufficient for a replication experiment. Strobe and Strack, 2014 believe conceptual replications are advantageous as "failures of exact replications do not tell us why findings cannot be replicated, they are ultimately not very informative" (2014, 64). Leonelli, 2018 argues that direct replications are non-essential and claims that non-replicable research also has epistemic value. Feest, 2019 claims that the replication crisis is exaggerated and improvement in science does not rely on the need for research to be validated to such extent. Machery, 2019 believes embodying a "Resampling Account", in which the replication Experiment A resamples the experimental components that have been treated as random factors of the original Experiment B, will allow scientists to dissolve the direct-conceptual replication distinction and refrain from assigning superiority to any one of those concepts. Lastly, some philosophers claim that even if direct replication is possible, the true validity of an experiment lies in confirmatory power (Collins, 1985). As new differences are introduced in an experiment during its replication process, the confirmatory power increases such that generalizability is possible. Collins implies that Experiment A does not need to identically replicate Experiment B and should have valid differences.

Like Collins, 1985, the variety of evidence thesis is accepted by philosophers that believe that ceteris paribus varied evidence has greater confirmatory power than experiments with less variation. Machery, 2019 explains how robust papers in psychology are trustworthy because of the variety of conceptual replications that are

employed by researchers. VET supports the notion of conceptual replication; however, it cannot be ignored that VET itself needs further investigation as inadequate research practices occur whilst performing conceptual replications. Psychological research papers that were supported by Machery, 2019, have engaged in wrongful practices including conducting conceptual replications with publication biases, confirmation biases, and lower statistical power. In this example, direct replication has a greater epistemic value (Shimmack, 2012). Consumer credibility in science is gained by peer review; however, scholars of peer-review have listed a variety of biases that inevitably enter the reviewing process including gender, language, nationality, reviewer conservatism, and confirmation bias (Longino, 2015). Despite VET being a promising solution, it must correctly identify when direct versus conceptual replication should be utilized.

Lastly, the replication crisis questions the nature of scientific error. The self-corrective thesis claims that science is epistemically positioned to correct its errors in future experiments. If we believe that scientists have engaged in non-replicable experiments and trust the claims of low replicability, then persisting false positives in publication will never be corrected. Romero, 2019 believes that the SCT is a borderline between the theoretical and practical aspects of science. The gap between the replicability crisis and our view of science stems from the ideal of SCT and reality (Longino, 2015). This gap may be a product of the replication crisis being normatively inadequate. On the other hand, there exists the plausibility that the ideal is adequate, and the gap is the result of scientists not doing their jobs correctly or systemic errors. Thus, it becomes the work of philosophers to conduct social epistemological work to fill this gap and understand the root of the SCT conflict.//

3. CONCLUSION

The replication crisis persists in social, behavioral, and biomedical sciences. Philosophers have suggested various statistical reforms to improve the replication crisis

including altering how frequentist statistics are completed, advocating descriptive statistics instead of frequentist statistics, or using Bayesian inferences. Benjamin et. al, 2018 have suggested changing the standard p-value from/p<0.05 to stricter/p<0.005. Similarly, Stegenga, 2020 claims that p-hacking is epistemically pernicious and can be battled via predictivism, Bayesian confirmation theory, and model selection theory. A methodological reform that I highly support advocates the notion of "*open*/*science*" in which researchers increase the transparency of their research practices by openly sharing their experimental designs, software, raw data to the scientific community. As philosophers lay out a foundation for what constitutes "*replication*", *when* to employ direct versus conceptual replications, and what constitutes a successful versus an unsuccessful replication, it becomes clear on how to approach the replication crisis. Needless to say, "When well-conducted replications are successful, they can provide us with greater confidence about the veracity of the predicted effort." (Brandt, 2014)

4. ACKNOWLEDGEMENTS

I thank the referee for reviewing my paper.

5. REFERENCES

[1] **Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., Johnson, V. E. (2018).** Redefine statistical significance./*Nature Human Behaviour*,/*2*(1), 6–10.

[2] **Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., … van't Veer, A. (2014).** The Replication Recipe: What makes for a convincing replication?/*Journal of Experimental Social Psychology*,/*50*, 217–224.

[3] **Collins, H. M. (1985).** *Changing order: Replication and induction in scientific practice*. London, England: SAGE Publications.

[4] **Feest, U. (2019).** Why replication is overrated./*Philosophy of Science*, (705451). doi:10.1086/705451

[5] **Hitzig, Z., & Stegenga, J. (2020).** The problem of new evidence: P-hacking and pre-analysis plans.,*Diametros*,*17*(66), 1–24.

[6] **Hull, D. L. (1995).** The golem: What everyone should know about science. Harry Collins, Trevor pinch./*Philosophy of Science*,/*62*(3), 487–488.

[7] **Hutchison, T. W., & Popper, K. R. (1959).** The logic of scientific discovery./ *Economica*,/*26*(103), 262.

[8] **Kerr, N. L. (1998).** HARKing: hypothesizing after the results are known./ *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc*,/*2*(3), 196–217.

[9] **LeBel, E. P., Berger, D., Campbell, L., & Loving, T. J. (2017)**. Falsifiability is not optional./*Journal of Personality and Social Psychology*,/*113*(2), 254–261.

[10] **Leonelli, S. (2018).** Rethinking reproducibility as a criterion for research quality. In/*Including a Symposium on Mary Morgan: Curiosity, Imagination, and Surprise*/(pp. 129–146). Emerald Publishing Limited.

[11] **Levin, N., & Leonelli, S. (2017).** How does one "Open" Science? Questions of value in biological research./*Science, Technology & Human Values*,/*42*(2), 280–305.

[12] **Longino, H. E. (2015).** The social life of behavioral science. In/*Studying Human Behavior*/(pp. 179–202). University of Chicago Press.

[13] **Longino, H.E. (2015)** The social dimensions of scientific knowledge. In E. N. Zalta (Ed.), The Stanford encyclopedia of philosophy (Spring 2015 ed.). Stanford, CA: Stanford University.

[14] **Machery, E. (2019).** What is a replication? doi:10.31234/osf.io/8x7yn

[15] **Resnick, B. (2018, June 6).** The "marshmallow test" said patience was a key to success. A new replication tells us s'more. Retrieved March 28, 2021, from Vox website:ttps://www.vox.com/science-and-health/2018/6/6/17413000/

marshmallow-test- replication-mischel-psychology

[16] **Romero, F. (2019a).** Philosophy of science and the replicability crisis./ *Philosophy Compass*, *14*(11). doi:10.1111/phc3.12633

[17] **Romero, F. (2019b).** Philosophy of science and the replicability crisis.

[18] **Schimmack, U. (2012).** The ironic effect of significant results on the credibility of multiple-study articles./*Psychological Methods*,/*17*(4), 551–566.

[19] **Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011).** False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant: Undisclosed flexibility in data collection and analysis allows presenting anything as significant./*Psychological Science*,/*22*(11), 1359–1366.

[20] **Schmidt, S. (2009).** Shall we really do it again? The powerful concept of replication is neglected in the social sciences./*Review of General Psychology: Journal of Division 1, of the American Psychological Association*,/*13*(2), 90–100.

[21] **Stroebe, W., & Strack, F. (2014).** The alleged crisis and the illusion of exact replication./*Perspectives on Psychological Science: A Journal of the Association for Psychological Science*,/*9*(1), 59–71.

RESEARCH ARTICLE

ANALYSIS OF CES PRODUCTION FUNCTION MODEL UNDER NON-NEUTRAL TECHNOLOGICAL PROGRESS FOR INDUSTRIAL SECTOR OF INDIA

M. K. Dave[1] and S. G. Raval[2]

ABSTRACT

Technological progress or change is the discovery of the new and improved technique of production. Technological change brings about an increased output per unit of labour. This paper analyses Constant Elasticity of Substitution production function model under non-neutral technological progress for industrial sector of all India for the period 1981-2017. The model seems to be fitted well for the data used at current prices as well as at constant prices.

KEYWORDS:

NON-NEUTRAL CES, REGRESSSION, CONSTANT AND CURRENT PRICES

1. INTRODUCTION

Production function model is characterized by its mathematical formulation, thus establishing a relationship between the output in terms of various factors of production employed. Most widely used and important production functions are Cobb-Douglas and CES production functions. Cobb-Douglas production function assumes that elasticity of substitution of input factors of production is unity, whereas CES production function assumes constant elasticity of substitution. CES production function has strong theoretical improvement over Cobb-Douglas production function.

(1) Research Scholar, Statistics Dept., Gujarat University, Ahmedabad, India. email : maheshkdave028@gmail.com

(2) Head of Statistics Department, Som Lalit Commerce College, Ahmedabad, Gujarat, India. email : drsgraval@gmail.com

(rcd. May '22 / rvd. June '22)

Technological change brings improved methods of production and shows changes in production techniques. Therefore an analysis pertaining to technological change is necessary. Technological progress is classified into mainly two categories-neutral and non-neutral by Murray Brown (1966). If at a constant capital-labor ratio, marginal rate of technical substitution is constant then technological progress is neutral. If marginal rate of technical substitution declines at a constant capital-labor ratio, technological progress is capital-deepening. Technological progress is labor-deepening if at a constant capital-labor ratio marginal rate of technical substitution increases.

In this paper , we have tried to analyse CES production function model under non-neutral technological change for time series data of all India for all industries. Section -2 discusses the data base. Section 3.1 gives model specification and estimation procedure. Section-4 shows statistical analysis by means of table 4.1 to 4.4. Concluding remarks are discussed in section-5.

A graphical presentation of the observed and estimated values for current as well as constant prices is given to visualize the growth pattern for the industries. This is given in Appendix-I at the end.

2. DATA BASE

For our study, we have used the data from Annual Survey of Industries (ASI) and Census of Indian Manufacturers(CIM). The data are for all India pertaining to all industries for the period 1981-82 to 2017-18. We have also considered data for Wholesale price index - Annual average for the period 1981-82 to 2017-18 taking 1981-82 as base year from RBI bulletin, applying forward & backward splicing method to obtain series at constant prices.

3. RESEARCH METHODOLOGY

3.1 MODEL

Let us consider the general form of non-neutral production function defined by Murry Brown for CES production function as under

$$Q_t = Ae^{\gamma t}\left[\alpha K_t^{-\rho} + (1-\alpha)L_t^{-\rho}\right]^{-\lambda/\rho} \cdot e^{u_t},\ 0 < \alpha < 1 \qquad (1)$$

where $A(t) = Ae^{\gamma t}$ which represents non-neutral technological change and

Q_t = Output for the year t

K_t = Working Capital for the year t

L_t = Wages to workers for the year t

γ= Technology parameter

α= Distribution parameter

ρ= Substitution parameter

λ= Degree of homogeneity

u_t= Disturbance term

Logarithmic form of equation (1) can be given by

$$L_nQ_t = L_nA + \gamma t - \frac{\lambda}{\rho}L_n\left[\alpha K_t^{-\rho} + (1-\alpha)L_t^{-\rho}\right] + u_t \quad (2)$$

The above equation (2) is not in linear or log-linear form .So using Taylor's series expansion around $\rho = 0$ and up to second order approximation, the expansion leads to

$$L_nQ_t = L_nA + \gamma t + \lambda L_nL_t + \lambda\alpha L_n\left(\frac{K_t}{L_t}\right) - \frac{\lambda\rho\alpha(1-\alpha)}{2}\left[L_n\left(\frac{K_t}{L_t}\right)\right]^2 + u_t \quad (3)$$

and therefore

$$L_nQ_t = \beta_0 + \beta_1 t + \beta_2 L_nL_t + \beta_3 L_n\left(\frac{K_t}{L_t}\right) + \beta_4\left[L_n\left(\frac{K_t}{L_t}\right)\right]^2 + u_t \quad (4)$$

Where $\beta_0 = Log_eA \implies A = e^{\beta_0}$

$\beta_1 = \gamma$

$\beta_2 = \lambda$

$\beta_3 = \lambda\alpha \Rightarrow \alpha = \frac{\beta_3}{\beta_2}$

$\beta_4 = \frac{\lambda\rho\alpha(1-\alpha)}{2} \implies \rho = -\left[\frac{2\beta_4\beta_2}{\beta_3(\beta_2-\beta_3)}\right]$

Equation (4) can be estimated by GLS method which gives estimates of $\beta_0, \beta_1, \beta_2$, β_3 and β_4 and hence estimates of γ, λ, α and ρ can be obtained.

4. STATISTICAL ANALYSIS

Regression Analysis for all Industries of All India (Current Prices)

Model: $L_nQ_t = \beta_0 + \beta_1 t + \beta_2 L_nL_t + \beta_3 L_n\left(\frac{K_t}{L_t}\right) + \beta_4\left[L_n\left(\frac{K_t}{L_t}\right)\right]^2$

Table: 4.1

Current prices						
Multiple R	0.99906948					
R Square	0.99813983					
Adjusted R Square	0.99790731					
Standard Error	0.06794228					
Observations	37					
ANOVA						
Regression	4	79.2628152	19.8157	4292.688	3.48755E-43	
Residual	32	0.1477169	0.004616			
Total	36	79.4105321				
Intercept	10.3692248	0.84100881	12.32951	1.07E-13	8.656145961	12.0823037
t	0.09324415	0.00674413	13.82597	4.81E-15	0.079506803	0.10698149
ln lt	0.36959789	0.06831346	5.410323	6.02E-06	0.230447925	0.50874786
ln k\l	0.5936225	0.15722307	3.77567	0.000655	0.273369591	0.91387541
[ln k\l]sq	-0.0393428	0.04017918	-0.97918	0.334835	-0.1211850	0.04249953

Table: 4.2

Parameter	A	γ	λ	α	ρ	σ	DW
Value	31862.98	0.09324	0.3692	0.606	-0.2189	1.28	1.88

Table 4.1 shows that at current prices all regression coefficients are found to be significant for the industrial sector of all India. The values of R^2 and $\bar{R}^2$ are also highly significant at 5% level of significance. About 99% variation is explained by the fitted model at current prices.

Table 4.2 gives estimated values of the parameters of fitted model at current prices. The technological parameter γ has the value 0.093 which indicates that technological change is almost negligible in the industrial sector of India. The estimated value of λ is 0.3996 for industrial sector of all India. The estimated value of substitution parameter ρ is -0.2189 and therefore the value of elasticity of

substitution is 1.28 which shows that the amount of substitution between input factors is reasonably good. The value of $\hat{\alpha}$ is 0.60. DW is 1.88 which indicates absence of auto correlation.

Regression Analysis for all Industries of All India (Constant Prices)

Model: $L_n Q_t = \beta_0 + \beta_1 t + \beta_2 L_n L_t + \beta_3 L_n \left(\frac{K_t}{L_t}\right) + \beta_4 \left[L_n \left(\frac{K_t}{L_t}\right)\right]^2$

Table: 4.3

Constant prices						
Multiple R	0.997794103					
R Square	0.995593072					
Adjusted R Square	0.995042206					
Standard Error	0.056942453					
Observations	37					
ANOVA						
Regression	4	23.44057532	5.860144	1807.324	3.42625E-37	
Residual	32	0.103758173	0.003242			
Total	36	23.54433349				
Intercept	7.010143342	0.526579652	13.3126	1.36E-14	5.937535697	8.082751
t	0.056541564	0.00236246	23.93335	5.42E-22	0.051729392	0.0613537
ln It	0.434775183	0.055997934	7.764129	7.45E-09	0.320711124	0.5488392
ln K/L	0.535414998	0.179554832	2.981902	0.00544	0.169673774	0.9011562
ln K/L sq	-0.093293	0.0401704	-2.32244	0.026728	-0.1751179	-0.011469

Table: 4.4

Parameter	A	γ	λ	α	ρ	σ	DW
Value	1107.8	0.05654	0.4348	0.2313	-1.5063	-1.975	1.95

Table 4.3 describes the estimated values of all regression coefficients for CES

production function model under non-neutral technical change. Here all coefficients $\beta_0, \beta_1, \beta_2, \beta_3$ are significant at 5% level of significance except coefficient β_4 for industrial sector of India at constant prices. The values of R^2 and $\bar{R}^2$ for the fitted model are 0.9955 and 0.9950 respectively which are highly significant values. About 99% of variation is explained by the fitted model.

From Table 4.4 it can be observed that the degree of homogeneity λ is 0.434 for industrial sector of all India at constant prices. The estimated value of substitution parameter ρ is -1.5063 which is insignificant. The estimated value of distribution parameter α is 0.2313. At constant prices also technological change is negligible for the industrial sector of India, as the estimated value of technological parameter γ is 0.057. DW is 1.95 which indicates absence of auto correlation.

5. CONCLUDING REMARKS

In this paper we have made a humble effort for visualizing the pattern of growth in the entire industrial sector of India by means of the fitted models of CES production function with non-neutral technological change at current as well as at constant prices.

As viewed from our analysis carried out at current as well as at constant prices, what we observed is that production function model posed for the industrial sector is found to be relevant to represent the production growth pattern during the period of our study. The statistical analysis is carried out with the help of **EXCEL** and **R** packages. The limitation of this application may be attributed due to the relevant limitations in the statistical data base. This analysis may be considered to be fruitful and important for the industrial sector of all India during this period of study. It can be used to visualize the economic growth pattern for all industries with their due limitations. The pattern for substitution effect under non-neutral technological change may be considered to be significant as viewed from our analysis. Such an analysis can also focus on the overall growth pattern for the industrial sector of all India. This can also be visualized for fruitful government planning exercises.

6. ACKNOWLEDGEMENT

We thank the referee for the comments which have helped us in thoroughly revising the earlier draft of this paper.

7. SELECTED REFERENCES

[1] **Allen,R.G.D.** (1938),Mathematical Analysis for Economist, Macmillan, London.

[2] **Arrows, K.J. Chenery H.B., Minhas, B.S. and Solow, R.M. (1961**), Capital-labour substitution and economic efficiency, *Review of Economic and Statistics*, Vol.43, No. 3 pp.225-250

[3] **Carlos Pestana Barros** (2008), Technical change and productivity growth in airports : A case study, Technical University of Lisbon, Portugal.

[4] **Damodar Gujarati and Sangeetha (2007),** Basic Econometrics, 4th ed. (Mcgraw Hill Co.)

[5] **Ferguson C.E.** (1965), Substitution Technical progress and Return to scale, *American Economic Review*, Vol. 55.

[6] **K. C. Soni (1988),** Productivity Trends in Production Functions with application to India data, (Ph.D. Thesis, Gujarat University).

[7] **Kmeneta, J. (1967),** On the Estimation of the CES Production Function, *International Economic Review*, Vol. 8, pp. 180-189.

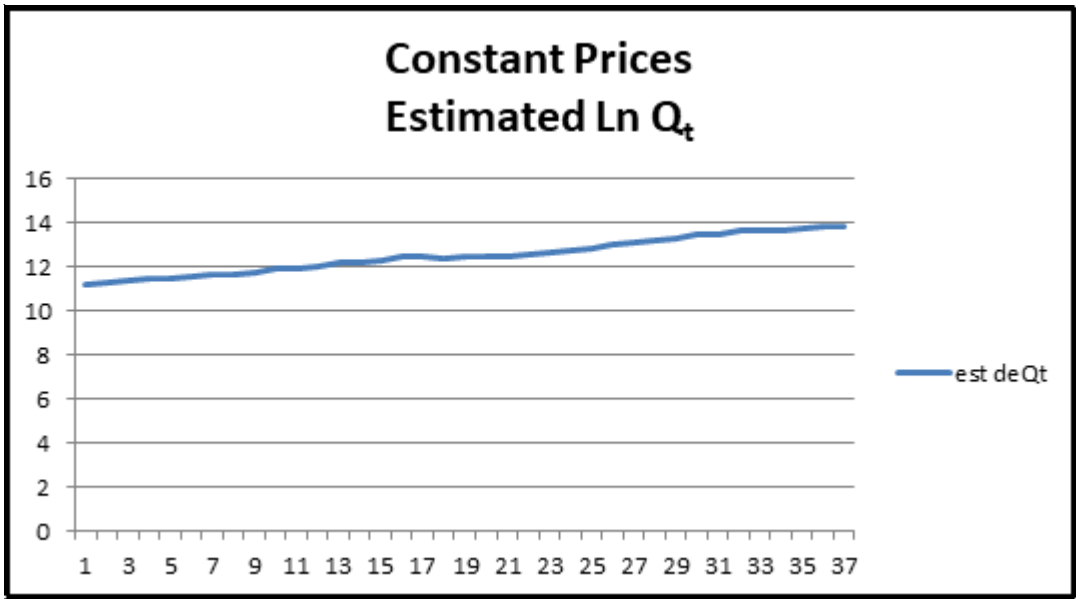
[8] **M. B. Thakar (1999),** Some Studies in Production functions with Application to India data, (Ph.D. Thesis, Gujarat University).

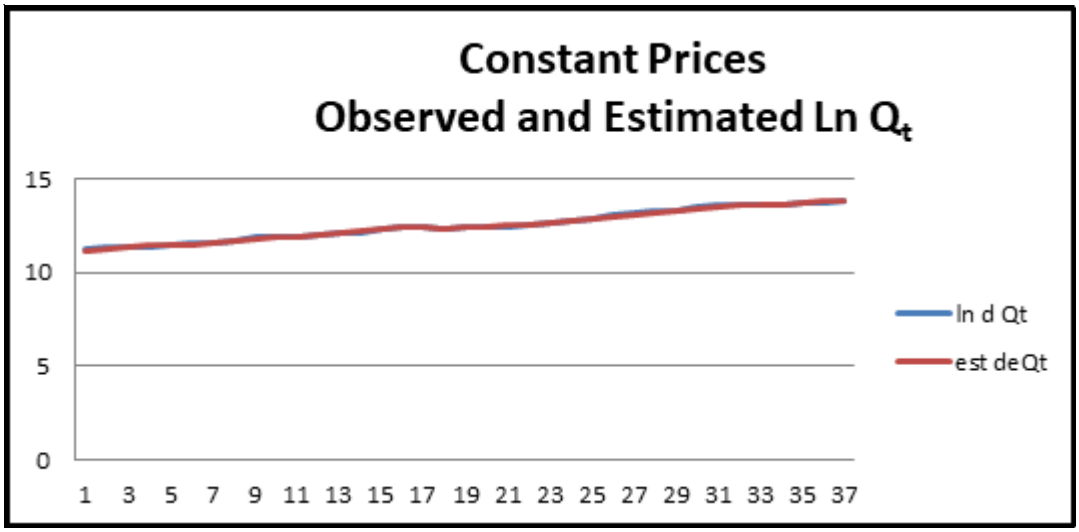
[9] **Murray-Brown** (1966), On the Theory and Measurement of Technical change, Cambridge University Press.

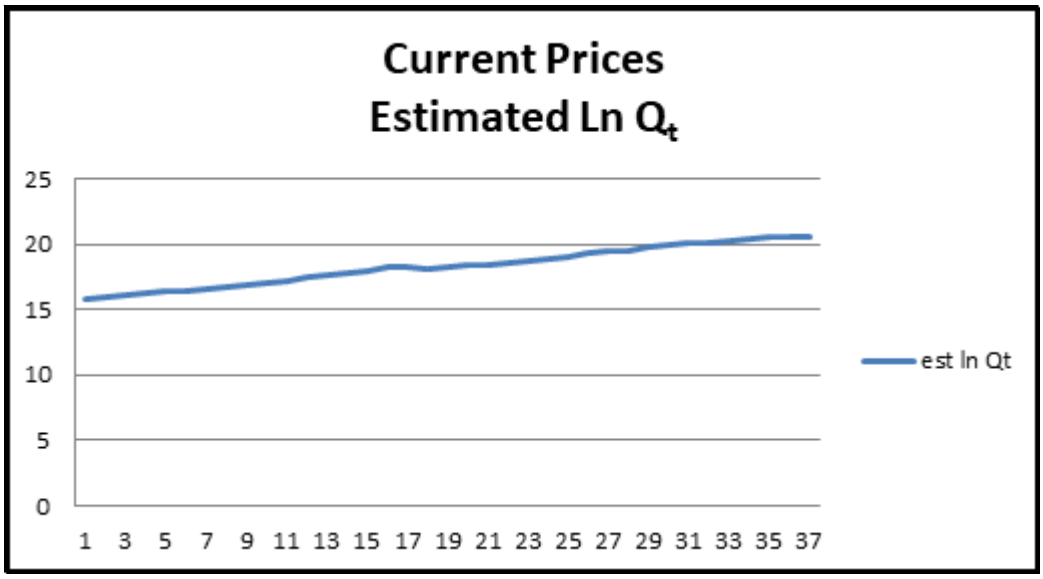
[10] **Thursby J. (1980),** Alternative CES Estimation Techniques, *Review of Economics and Statistics*, Vol. 62(2), pp. 295-299.

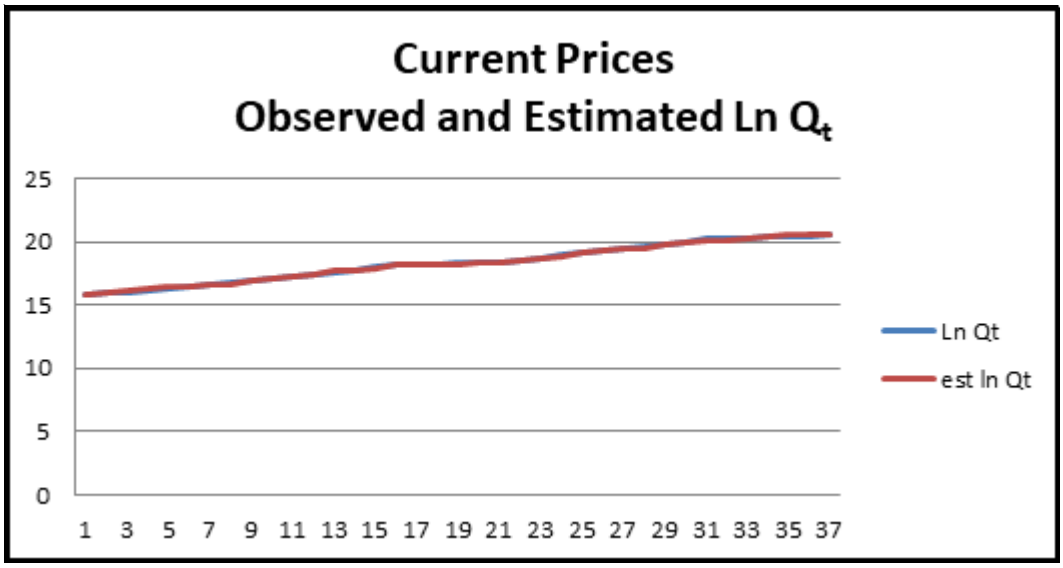
[11] **Walters A. A. (1963),** Production and Cost functions, An Econometric Survey, *Econometrica* Vol. 31(1, 2).

9. APPENDIX-I: GRAPHICAL PRESENTATION






Current Prices
Estimated Ln Q_t
25
20
15
10
5
0
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37
est ln Qt


Current Prices
Observed and Estimated Ln Q_t
25
20
15
10
5
0
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37
Ln Qt
est ln Qt

RESEARCH ARTICLE

A STUDY RELATED TO THE IMPACT OF CORONA IN TRICHY DISTRICT

P. Mariappan[1] and V. A. Annie Preethi[2]

ABSTRACT

The purpose of this study is the assessment of the impact of corona pandemic in Trichy District. Scope of this work is to investigate the effects of Corona, current events and assessment through the interpretation of how the city will recover after Corona and how it can be sustainable. The present study helps to analyze the future with few measures and speedy recovery rate. Also the study has discovered the impacts of the outbreaks by using the statistical models.

KEYWORDS

Impacts, current events, measures, outbreaks, statistical models.

INTRODUCTION

Corona Virus or COVID-19 originated in Wuhan, China in the month of December 2019. Till date, there is no avowed human immunization for combating it. COVID-19 has affected our lives to a great extent. Work, economy, education and almost everything has come to stand still. People nowadays only focus on what is essential to live. The reasons for its spread and thinking about its threat, practically all the nations have proclaimed either partial or complete lockdowns all through the influenced districts and regions. Since there is no endorsed medicine till now for

(1) Head and Associate Professor. P. G. and Research Department of Maths. Bishop Heber College, TRICHY, Tamilnadu, India. email : mathmari@yahoo.com

(2) P. G. Student, P.G. and Research Dept. of Maths, Bishop Heber College, TRICHI, Tamilnadu, India. email : ma201207002@bhc.edu

(rcd. March '22 / rvd. June '22)

slaughtering the infection so the legislatures of all nations are concentrating on the precautionary measures which can stop the spread. This virus causes respiratory tract infections that can range from mild to fatal. This Infection can influence individuals of any age yet older people or individuals with previous ailments are more inclined or vulnerable to it. The most testing part of its spread is that an individual can have the infection for a long time without indicating side effects [3–6]. Utilization of mask, sanitizer, customary hand washing and cleanliness is the most ideal approach to keep from this illness. People are urged to believe Staying homes safes lives. COVID-19 will reshape our reality.

Machine learning plays a major role in better understanding and examining COVID-19 crisis as it identifies the patterns in data and uses them to automatically make predictions or decisions. On the basis of this concept, this work has been carried out for analysis of COVID-19 cases and prediction of upcoming new positive cases.

RELATED TO WORK

Since the time the rise of COVID-19 and its resulting spread across landmasses overwhelming both progressed and creating countries, there has been a great deal of exploration papers distributions on different parts of COVID-19. So, in various research paper analysis is done on vaccination, drug therapy and also on the prediction of future infected, recovered and death cases. Using various forecasting techniques L. Jia et al., 2020 did the prediction and analysis of COVID19 through various models.

According to this paper they concluded that this epidemic will be over presumably in the last of April but their estimation is not so accurate. F. Rustam et al. (2020) introduced future forecasting on Covid-19 using various machine learning models. They use four regression models namely LR, LASSO, SVM and Exponential Smoothing. They predicted their results on the basis of Evaluation Parameters used in their models. According to their results ES model is best for predicting infected

cases and recovered cases and LR model best predict the death rate.

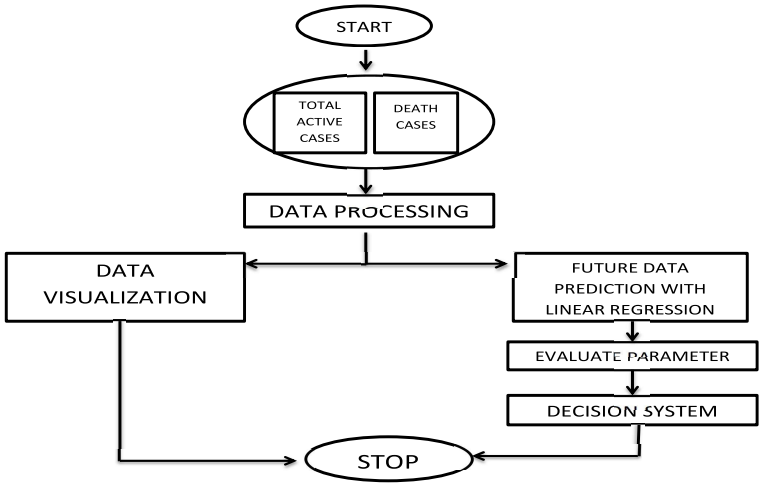
MOTIVATION

As we know, data science helps us to clarify and understand the data that has been accumulated so far. Data science helps one to simulate and imagine the patterns of how coronavirus spreads including the number of patients reported monthly with coronavirus and the death cases, i.e. it is growing or more than before and so on. Data analysis helps one to gain some valueable insight into the data. So, in our model, we first try to imagine data and then forecast future data and also try to find the best fit regression model that could help us with future predictions.

MATERIALS AND METHODS

In this examination paper we have utilized the overall details of coronavirus from August, 2020, to November, 2021, was assembled from the online resources from Covid-19 Bulletin inTiruchirapalli corporation website. The datasets give us the quantity of affirmed cases and death cases throughout the Tiruchirapalli District. And in this paper Linear Regression Model is used.

EXPERIMENT WORK FLOW CHART



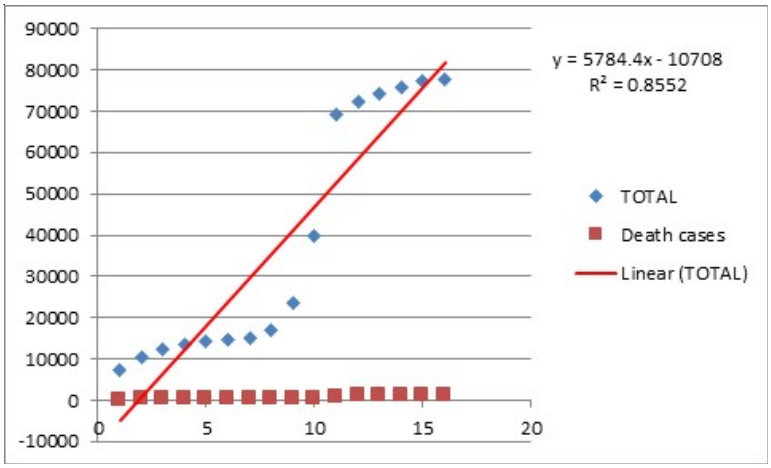
Tabulation:

Let dependent variable be Death cases.

Let independent variable be Total cases.

Month	Total cases	Death cases
August	7478	120
September	10527	145
October	12505	168
November	13421	172
December	14183	176
January	14646	180
February	14984	183
March	17134	190
April	23366	211
May	39803	368
June	69244	910
July	72422	964
August	74354	1007
September	75990	1029
October	77497	1059
November	77789	1061

This is the graphical representation of the data which we collected. It is drawn before the model was applied.



Linear Regression Equation:

Mathematically, a linear regression is defined by this equation

$y = bx + a + \varepsilon$, where

- **x** – is an independent variable,
- **y** – is an dependent variable,
- **a** – is the Y- intercept, which is the expected mean value of y when all x variables are equal to 0,
- **b** – is the slope of regression line, which is the rate of change for y as x changes,
- **ε** – is the random error term, which is the difference between the actual value of a dependent variable and its predicted value

In statistical modeling, **Regression analysis** is used to estimate the relationships between two or more variables. **Dependent variable** (*criterion* variable) is the main factor you are trying to understand and predict. **Independent variables** (*explanatory* variables, or *predictors*) are the factors that might influence the dependent variable. Regression analysis helps you understand how the dependent variable changes when one of the independent variables varies and allows to mathematically determine which of those variables really has an impact.

Regression Analysis Output:

Regression Statistics	
Multiple R	0.993123931
R Square	0.986295143
Adjusted R Square	0.985316224
Standard Error	3608.459238
Observations	16

Multiple R – it is the Correlation coefficient that measures the strength of a linear relationship between two variables. The Correlation coefficient can be any value from **-1 to 1**, and it is the absolute value indicates the strength of the regression.

So, here we have 0.99 value and it is strong positive relationship.

R square – it is the Coefficient 0f determination, which is used as an indicator of the goodness of fit. It shows how many points fall on the regression line. The R^2 is 0.98, which is fairly good. It means that 98% of our values fit the regression model.

Adjusted R square – It is the R square adjusted for the number of independent variable in the model.

Standard Error – it is an absolute measure that shows the average distance that the data points fall from the regression line.

Regression analysis output : Anova

ANOVA					
	df	***SS***	***MS***	***F***	***Significance F***
Regression	1	13119099156	13119099156	1007.535616	1.91368E-14
Residual	14	182293693	13020978.07		
Total	15	13301392849			

- **df -** is the number of the degrees of freedom associated with the sources of variances.
- **SS –** is the sum of squares,
- **MS –** is the mean square,
- **F –** is the F statistic, it is used to test the overall significance of the model,
- **Significance F** is the P-value of F.

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	2784.006064	1441.177647	1.931757733	0.073886602	-307.0125673	5875.025	-307.013	5875.025
Death cases	71.86187876	2.263957989	31.74170152	1.91368E-14	67.0061718	76.71759	67.00617	76.71759

The most useful component in this section is **Coefficients**. It enables you to build a linear regression equation

Regression formula : y = bx + a

- y = 71.86187876 x + 2784.006064

a (y – intercept) : 2784.006064

b (slope of a regression) : 71.86187876

Correlation Coefficient : 0.993123931

The most useful component in this paper is coefficients. It enables you to build a Linear Regression : **y = 71.8618 x + 2784.**

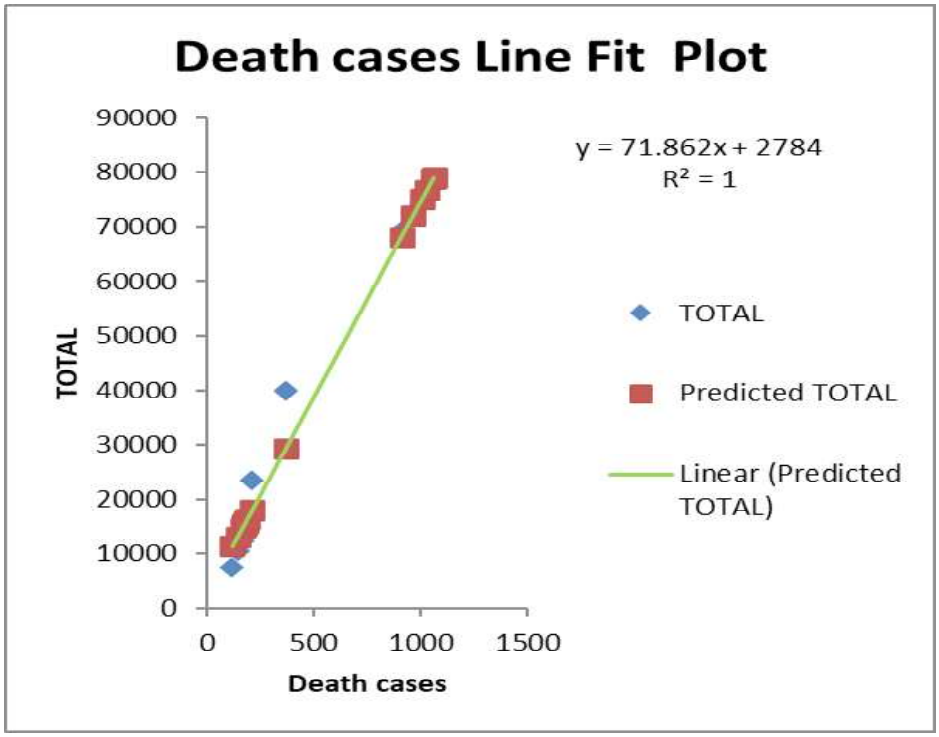
RESIDUALS

The independent variables are never perfect predictors of the dependent variables. And the residuals can help you to understand how far away the actual values are from the predicted values.

Observation	*Predicted TOTAL*	*Residuals*
1	11407.43152	-3929.43152
2	13203.97848	-2676.97848
3	14856.8017	-2351.8017
4	15144.24921	-1723.24921
5	15431.69673	-1248.69673
6	15719.14424	-1073.14424
7	15934.72988	-950.729877
8	16437.76303	696.2369718
9	17946.86248	5419.137518
10	29229.17745	10573.82255
11	68178.31573	1065.684266
12	72058.85719	363.1428133
13	75148.91797	-794.917973
14	76729.87931	-739.879306
15	78885.73567	-1388.73567
16	79029.45943	-1240.45943

However, our predicted values are high and our regression's significant value is also significant. The residual approximately 38458.

We have the graphical representation of the Linear Regression, where we clearly notice the predicted values of Total cases are high.



CONCLUSION

This examination introduced current patterns of COVID-19 occurrence from August 2020 to November 2021 as envisioned in our task. The quickly expanding number of latest COVID-19 cases day by day worldwide has placed an overwhelming weight on clinical assets in nations with enormous flare-ups. Therefore, prediction of future confirmed cases became necessary. We have done prediction by using linear regression model so that we can conclude which model is best and on the basis of that model we are able to tell the rate of increase in number of infected, and death cases in future.

ACKNOWLEDGEMENT

We thank the referee for reviewing our paper.

REFERENCES

[1] **Petropoulos F, Makridakis S, Braunstein LA.** Forecasting the novel coronavirus COVID-19. PLoS ONE 2020;15(3):e023123doi:https://doi.org/10.1371/journal.pone.023123610.1371/journal.pone.0231236.g00110.1371/journal.pone.0231236.g00210.1371/journal.pone.0231236.g003.

[2] **Al-qaness MAA, Ewees AA, Fan H, Abd El Aziz M.** Optimization method for forecasting confirmed cases of COVID-19 in China. J Clin Med 2020;9(3):674. doi: https://doi.org/10.3390/jcm9030674.

[3] **Fanelli D, Piazza F.** Analysis and forecast of COVID-19 spreading in China, Italy and France. Chaos Solitons Fractals 2020;134:109761. doi: https://doi.org/10.1016/j.chaos.2020.109761.

[4] **Sperrin M, Grant SW, Peek N.** Prediction models for diagnosis and prognosis in Covid-19. BMJ 2020:m1464. doi:https://doi.org/10.1136/bmj.m1464. [5] Gupta R, Pandey G, Chaudhary P, Pal SK. SEIR and regression model based COVID-19 outbreak predictions in India; 2020. https://doi.org/10.1101/2020.04.01.20049825.

[6] **Perc M, GorišekMiksic´ N, Slavinec M, Stoz¡er A.** Forecasting COVID-19. Front Phys 2020;8. doi: https://doi.org/10.3389/fphy.2020.00127.

[7] **Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA,** et al. Realtime tracking of self-reported symptoms to predict potential COVID-19. Nat Med 2020;26(7):1037–40. doi: https://doi.org/10.1038/s41591-020-0916-2.

[8] **Hamzah FAB.** CoronaTracker: world-wide COVID-19 outbreak data analysis and prediction; 2020. https://doi.org/10.2471/blt.20.255695. [

9] **Rustam F, Reshi AA, Mehmood A, Ullah S, On B-W, Aslam W,** et al. COVID-19 future forecasting using supervised machine learning models. IEEE Access 2020;8:101489–99. doi: https://doi.org/10.1109/Access.628763910.1109/ACCESS.2020.2997311.

[10] **Sengupta S, Mugde S, Sharma G.** Covid-19 pandemic data analysis and forecasting using machine learning algorithms; 2020. https://doi.org/10.1101/2020.06.25.20140004

RESEARCH ARTICLE

A STATISTICAL STUDY FOR THE PERSONS WITH DISABILITIES IN GUJARAT STATE

Preeti D. Ravat*

ABSTRACT

This paper is based on the survey of persons with disabilities conducted during NSS 76th round (July – December 2018). The survey of person with disabilities provide the database regarding the incidence and prevalence of disability in the Gujarat along with various other indicators related to person with disabilities. The result of the survey are useful to various stakeholders like National Institution for Transforming India (NITI ayog), Ministry of social Justice and Empowerment, Ministry of Health and Family Welfare, different Institutes/Organizations, various researchers, etc.

KEY WORD

NSS Round, Disability, Literacy rate, Household size

1. INTRODUCTION:

The main objective of the Survey of person with disabilities conducted by NSO in its 76th round was to estimate indicators of incidence and prevalence of disability, cause of disability, age at onset of disability, facilities available to the persons with disability, difficulties faced by persons with disability in accessing/using public building and public transport, arrangement of regular care giver, out-of pocket expenses relating to disability, etc.

2. ALLOCATION OF SAMPLE:

The 76th round survey of persons with disabilities was spread over 406 FSUs from which 204 FSUs in rural area and 202 FSUs in urban area of Gujarat. Below

* Research Assistant, Directorate of Economics and Statistics, Gandhinagar.
M : 9824446756
(rcd. March '22 / rvd. June '22)

table shows the detailed allocation of FSUs and total number of households as well as number of disable person.

Table: 1 Allocation of FSUs, Number of households & persons surveyed & Number of persons with disabilities surveyed

	Number of FSUs allocated		Number of household surveyed		Number of persons surveyed		Number of persons with disabilities surveyed	
	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
Gujarat	204	202	2637	1747	13957	8131	2448	1451
All India	5384	3616	81004	37148	402589	173980	74946	31948

3. HOUSEHOLD SIZE AND PERCENTAGE OF DISABLE PERSON:

Table-2 presents average household size, sex ratio and percentage of persons with disabilities for male and female in rural and urban areas.

Table 2: Average household size, sex ratio and percentage of persons with disability

Indicators		Rural		Urban		Rural+Urban	
		Gujarat	All India	Gujarat	All India	Gujarat	All India
Average household size		4.9	4.6	3.8	3.8	4.4	4.3
Sex ratio (number of female per 1000 male)		939	938	848	908	899	929
Percentage of	Males with disability	1.8	2.6	1.6	2.1	1.7	2.4
	Females with disability	1.4	2.0	1.3	1.8	1.3	1.9
	Persons with disability	1.6	2.3	1.5	2.0	1.5	2.2

- In Gujarat prevalence of disability was 1.5%, among them 1.7% male and 1.3% female found disable.
- In Gujarat, sex ratio was higher in rural area which was 939 as compare to all India.

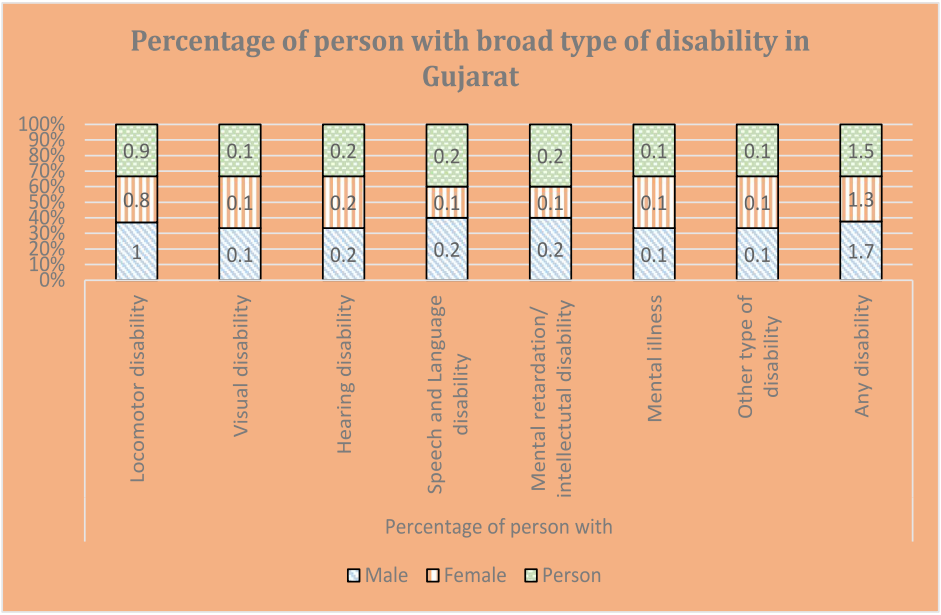
4. PERSON WITH BROAD TYPE OF DISABILITY

Table-3 presents the percentage of persons with disability like Locomotor, Visual, and Hearing etc. in Gujarat.

Table 3: Percentage of persons with broad type of disability in Gujarat (Rural + Urban)

Indicators		Male	Female	Person
Percentage of person with	Locomotor disability	1	0.8	0.9
	Visual disability	0.1	0.1	0.1
	Hearing disability	0.2	0.2	0.2
	Speech and Language disability	0.2	0.1	0.2
	Mental retardation/ intellectutal disability	0.2	0.1	0.2
	Mental illness	0.1	0.1	0.1
	Other type of disability	0.1	0.1	0.1
	Any disability	1.7	1.3	1.5

- It can be seen from above table that, in Gujarat most of the person suffering from locomotor disability. In rural Gujarat 0.9% of persons and in urban Gujarat 1% of persons were suffering from locomotor disability.



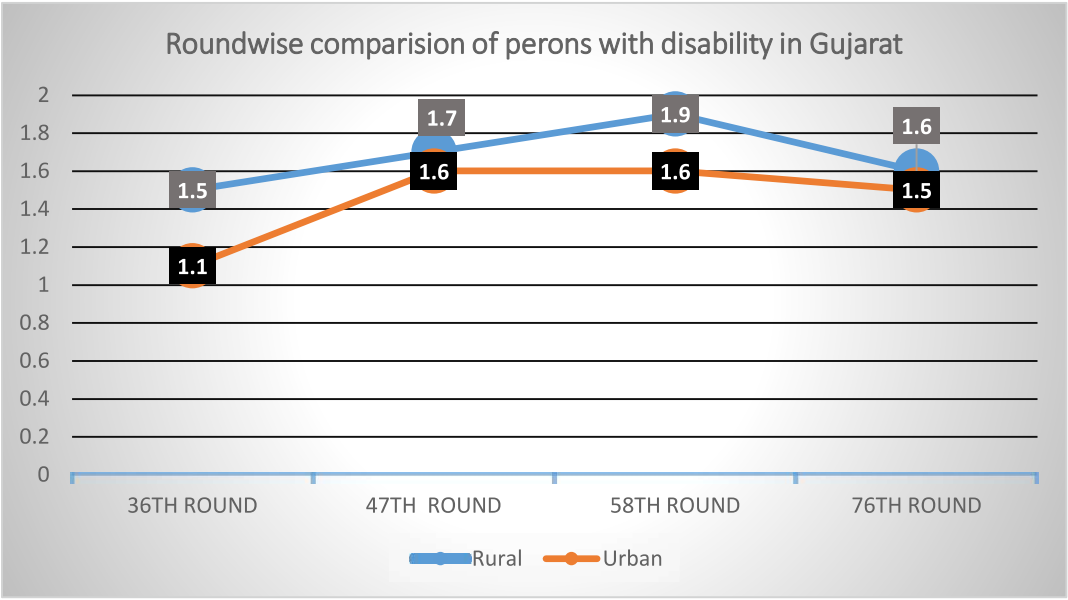
5. COMPARISON OF PREVIOUS ROUND OF NSS

Table 4: Percentage of person with disability in different round of NSS in Gujarat

Gender	NSS Round							
	36th Round (July-Dec.'81)		47th Round (July-Dec.'91)		58th Round (July-Dec.'02)		76th Round (July-Dec.'18)	
	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
Male	1.6	1.2	1.8	1.7	2.2	1.8	1.8	1.6
Female	1.4	1.0	1.6	1.6	1.6	1.3	1.4	1.3
Persons	1.5	1.1	1.7	1.6	1.9	1.6	1.6	1.5

- It was observed from above table that, in rural area as well as in urban area person with disabilities was lower in year 2018 as compared to year 2002.

Figure-2



6. LITERACY RATE OF DISABLE PERSON

Table 5: Literacy rate of disable person of age 7 years and above						
	Male		Female		Person	
	Gujarat	All-India	Gujarat	All-India	Gujarat	All-India
Rural	65.3	57.4	41.9	33.3	55.4	47.2
Urban	70.4	73.3	62.5	55.8	67.2	65.5
Rural + Urban	67.5	61.6	50.1	39.6	60.2	52.2

- Among persons with disabilities of age 7 years and above, 60.2% of persons in Gujarat and 52.2% of persons in India were literate.
- In Rural Gujarat, literacy rate of person with disability was higher in male with 65.3%, for all India it was 57.4%.
- In Urban Gujarat, literacy rate of person with disability was higher in male with 70.4%, for all India it was 73.3%.

7. DISABILITY SINCE BIRTH OR NOT

Table -6 presents the percentage of person who had disability since birth and had disability but not since birth.

Table 6: Disability since birth or not							
Indicator		Rural + Urban					
		Male		Female		Person	
		Gujarat	All India	Gujarat	All India	Gujarat	All India
Percentage of person	Who had disability since birth	38.9	29.2	33.6	27.2	36.7	28.3
	Who had disability but not since birth	60.6	69.8	65.8	72.0	62.8	70.7

- In Gujarat, 36.7% of persons who had disability since birth and 62.8% persons who had disability but not since birth.
- In Gujarat the ratio of disability since birth was higher for male with 38.9, where as the ratio of disability but not since birth was higher for female with 65.8%.

8. PERSONS RECEIVED HELP AND DISABILITY CERTIFICATE

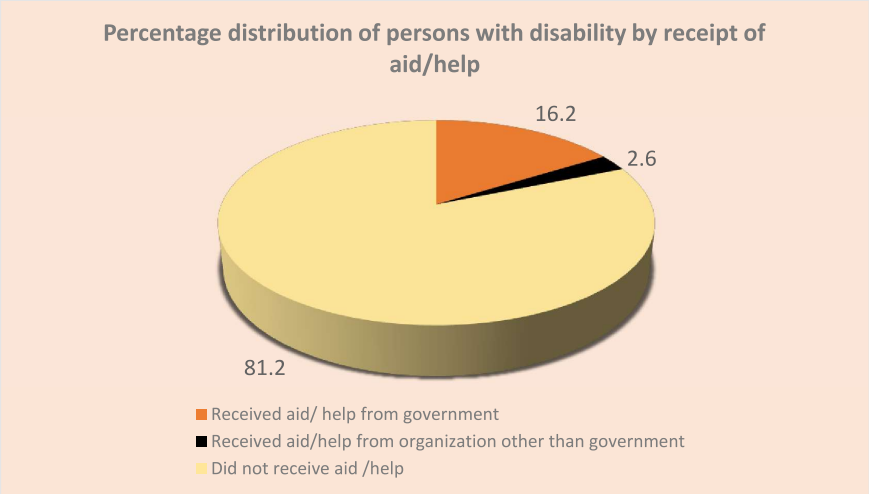
Table -7 presents the percentage distribution of persons with disability by received aid/help from government, Received aid/help from organization other than government, Did not receive aid /help and having certificate of disability.

Table 7: Percentage distribution of persons with disability by receipt of aid/help and having certificate in Gujarat

Indicator (Rural + Urban)		Male	Female	Person
Percentage distribution of persons with disability by receipt of aid/help	Received aid/ help from government	17.6	14.2	16.2
	Received aid/help from organization other than government	2.8	2.3	2.6
	Did not receive aid /help	79.6	83.5	81.2
	all	100	100	100
Percentage of persons with disability having certificate of disability		34.5	26.1	31.0

- In Gujarat, 31% of persons with disability who were having certificate of disability. 34.5% of male and 26.1% of female having certificate of disability.

Figure – 3



- 81.2% of persons in Gujarat who did not receive any aid/help while 16.2% of persons were received aid/help from government.
- Percentage of person who received aid/help from organization other than government were very less with 2.6%.

9. CONCLUSION AND SUGGESTION

- In Gujarat prevalence of disability was 1.5%, among them 1.7% male and 1.3% female found disable.
- It was observed that majority of the cases of locomotor disability were due to fractures in young and stroke in elderly people. Most of the fractures in young were due to road traffic accidents. Osteomyelitis, dislocation, osteoarthritis, rheumatoid arthritis, and cerebral palsy were the main causes, so it was must to develop the right attitude about driving.? Under age driving is illegal and we have to try awareness for that so young age accident has been prevent and decrease the percentage of locomotor disability.
- The substantial reductions in disability between the year 2002 and 2018 were likely due to advances in medical care as well as changes in socioeconomic factors.

- Among persons with disabilities of age 7 years and above, 60.2% of persons in Gujarat were literate. Different activities may also help the person with a disability to find something they are passionate about, which can rekindle their sense of enthusiasm for life.
- In the year 2017, emergency services 108 completed 10 years in Gujarat and it was the second state where this service was introduced. This was also reason behind low disability or deduction the percentage of disability as compare to previous year.
- Many people whose bodies are disabled but their mind still at the level of the knowledge, so engaging in courses or programs that stimulate their brain — as well as the creative flow of ideas and new learning.
- Establish companies that specialize in employing the disabled persons for paid work and will even organize transport. This may be a good option to look into if the person is still interested in working.
- In Gujarat, 31% of persons with disability who were having certificate of disability.

10. ACKNOWLEDGEMENT

I would like to thank Dr.Rakesh.R.Pandya sir, Director of DES, Gandhinagar and Shri S.S. Suthar sir, Joint Director of GSIDS for their useful suggestion to revise this research paper. I also thank the referee for reviewing my paper.

11. REFERENCES:

[1] Report on survey of Disabled Person –NSS 36th Round(July –Dec 1981)
[2] A Report on disabled person –NSS 47th Round (July Dec 199)-Report No. 393
[3] Disabled person in India-NSS 58th Round (July-Dec 2002)-Report No.485
[4] Person with disabilities in India –NSS 76th Round (July – Dec 2018)- Report No. 583

REVIEW ARTICLE

SOME MYTHS AND THEIR REMEDIAL IN AGRICULTURE EXPERIMENTAL DESIGN

D. K. Ghosh*

ABSTRACT

This article is a gist of invited talk delivered by me at Navsari Agricultural University, Navsari. In this article we have explained in detail along with suitable examples about the experimental design which will be suitable for the agricultural scientist to apply at the respective situations. Here we have discussed only the basic, factorial and confounding in factorial experiments.

KEW WORDS AND PHRASES:

experimental error, variability, coefficient of variations, generalized randomized block designs, and confounding.

1. INTRODUCTION

We always feel that in agricultural sciences, biological sciences, medical sciences, etc., designing an experiment is an inevitable component for each and every research endeavour. It is well known phenomena that the data generated through experimental design exhibit a lot of variability. The variability is namely (i) wanted and desirable; however, it is controllable in the sense that it can be accounted for, (ii) unwanted, undesirable and uncontrollable. But the reason for its presence is unknown. For example, consider plots as an experimental unit then this experimental units subjected to the same treatment give rise to different observations and thus create variability. However, these plots are expected to yield same response, but actually the responses

* UGC BSR Faculty Fellow, Saurashtra University, Rajkot, Gujarat, INDIA.
(rcd. Nov. 2021 / rvd. Dec 2021)

are different; however, the reasons are unknown. After conducting several experiments, it is observed that the variation may arise due to following variables like, fertility/ soil, moisture/soil depth, insecticide/soil depth, pesticide/soil area etc. of the plots. By using the theory of linear estimation and analysis of variance techniques, we can partition the total variability of the data into two major components. The first major component comprises of that part of the total variability in which we can assign causes or reasons. The second component contains that part of the total variability in which there is no possibility of assigning any cause or reason. Since we cannot assign any cause or reason to this part of variability and hence, this part is known as experimental error. Such variability arises because of some factors unidentified as a source of variation. An experimenter is always trying in conducting the planning of experiment carefully and systematically however, the planner is not sure of total elimination of this component. In any case we are interested to estimate the experimental error whenever, observations are obtained from experimental units.

In such situations, we should select a design that will minimize the experimental error as much as possible. However, no rule of thumb is available to explain what amount of experimental error is small and what amount of it can be termed as large. However, a popular measure of the experimental error is available which we called Coefficient of Variation (CV). In general, in any experiment, we feel if the CV is very small, then the experiment will be more precise, however, no degree of smallness is defined so far.

Sometimes, it happened that even a low value of CV (in percent) may be viewed with suspicion and hence, must be ascertained by viewing other parameters like variability across replications for all the treatments or otherwise. This shows that the explainable part of the total variability has two major components. Out of two components, treatments are one component which causes variability in every designed experiment. The other component of the explainable part of variability is the experimental error.

Experimental error is unwanted and undesirable. The factors that cause experimental errors are also called nuisance factors. This experimental error is eliminated using the principle of local control. Moreover, before planning and conducting an experiment, the researchers must have a complete knowledge about the experimental units (on which the experiment would be conducted), the sources of variation and the nature of variability of the experimental units. If this variability is substantial and is not accounted for by proper designing of experiment, then this component will merge with the experimental error and will be the reason so as to make experimental error unduly very large. This will provide us the least precise estimators. But a variability due to experimental units may be accounted for in several ways.

However, its remedy depends upon the sources of variations and nature of the factors causing variability in the experimental units. As a matter of fact, the least variability in the experimental units will be cause of recommending what types of design is to be applied in the corresponding situations.

We may select one blocking system, two blocking systems, nested blocks or nested rows and columns designs, etc. for eliminating the variability. However, depending upon practical constraints, a naive design is the best design. Treatments in the experimental design may be unstructured, i.e., treatments are the levels of a single factor. For example, four different varieties of paddy; three different combinations of fertilizers, etc. In such situation, we are interested to make all the possible pair wise comparisons among treatments. In case of factorial experiment, let treatments may be structured. In other words, we can say that treatments are all possible combinations of levels of several factors. For example, combinations of three levels of nitrogen, two levels of phosphorous, and three levels of potash, forming a total of 18 treatment combinations. In this case our interest is to estimate the main effects and interaction of the factors only. Once the data has been generated through the design, a model is to be decided. A typical model could be defined as

response = constant + explainable part of variability + unexplainable part of variability (or error)

Considering the fact that the explainable part of the variability has two major components viz., due to treatments and due to experimental units, the model could be modified as

response = constant + treatments effect + experimental units' effect + error.

Hence, it established that the model is defined by the way the experiment is designed and data are analysed. Let there is n observations.in an experiment.

The analysis of variance table is given in Table -1.

Table – 1 ANOVA table of unstructured design

Source of variations	DF	Sum of squares	Mean squares	F-ratio
Explainable part of Variability(A)	a	SSA	SSA/a = σ_A^2	σ_A^2/σ_E^2
Unexplainable part of variability, i.e., Error (E)	n – a -1	SSE	SSE/(n–a–1) = σ_E^2	
Total	n - 1	SST		

Further the analysis of variance can be restructured and is shown in Table–2.

Table – 2 Analysis of variance table of structured and unstructured treatments

Treatments Unstructured					Treatments Structured				
Sources of variations	df	Sum of squares	Mean squares	F-ratio	Sources of variations	df	Sum of squares	Mean squares	F-ratio
Treatments	a	SST	$SST/a = \sigma_T^2$	σ_T^2/σ_e^2	Treatments	a	SST	$SST/a = \sigma_T^2$	σ_T^2/σ_e^2
Experimental units	b	SSB	$SSB/b = \sigma_B^2$	σ_B^2/σ_e^2	Main effects	m	SSM	$SSM/m = \sigma_m^2$	σ_m^2/σ_e^2
Error	E	SSE	$SSE/E = \sigma_e^2$		All Interactions	a-m	SSI	$SSI/(a-m) = \sigma_I^2$	σ_I^2/σ_e^2
Total	n-1	SSTO			Experimental units	b	SSB	$SSB/b = \sigma_B^2$	σ_B^2/σ_e^2
					Error	E	SSE	$SSE/E = \sigma_e^2$	
					Total	n-1	SSTO		

Where, E = (n – a – b – 1)

From Table – 2, we can see that there is no difference in the analysis of variance

table, if the treatments are either unstructured or structured. We can also observe that if the treatments are structured, then the treatment sum of squares can be further partitioned into two sums of squares namely, (i) the sum of squares of main effects and (ii) sum of squares of interaction effects.

However, our main interest is to select a good design for a particular treatment setting. Following are the several experimental designs which we can select depending upon various situations.

1. COMPLETE RANDOMIZED DESIGNS

As the name implies, the completely randomized design (CRD) refers to the random assignment of experimental units to a set of treatments. It is essential to have more than one experimental unit per treatment to estimate the magnitude of experimental error and to make probability statements concerning treatment effects

A completely randomized design is a type of experimental design where the experimental units are randomly assigned to the different treatments. It is used when the experimental units are believed to be "uniform;" that is, when there is no uncontrolled factor in the experiment.

Completely randomized designs are the simplest in which the treatments are assigned to the experimental units completely at random. This allows every experimental unit, i.e., plot, animal, soil sample, etc., to have an equal probability of receiving a treatment.

Assumption of Complete randomized design

An assumption regarded to completely randomized design (CRD) is that the observation in each level of a factor will be independent of each other.

Difference between Randomized block design and completely randomized design

Randomized complete block designs differ from the completely randomized designs in that the experimental units are grouped into blocks according to known

or suspected variation.

Null Hypothesis

The null hypothesis for any ANOVA is always that there are no differences between any of the population means. That means that the alternative hypothesis must be that there are differences between some of the population means.

Suppose that $t_1, t_2, \ldots, t_t$ are t treatments. Further, suppose that each treatment is replicated r times. Since there are t treatments and each is replicated r times, so we need n = t x r experimental units, where n denotes total number of experimental units.

However, in case of completely randomized design treatment may be replicated an unequal number of times also. Suppose that $t_1, t_2, \ldots, t_t$ are replicated $r_1, r_2, \ldots, r_t$ times, respectively, where, $r_1 + r_2 + \ldots + r_t$ = n.

Now the complete experiment is divided into n number of experimental units. The experimental units are numbered serially, starting with one and ending with the total number of experimental units, n, in a serpentine manner. For example, there are six treatments, $t_1, t_2, t_3, t_4, t_5, and\ t_6$ and each is repeated five times. So, we need 30 experimental units. The thirty experimental units are numbered and is shown in Table – 3. Here we discuss the example only for equal number of replications.

Table – 3 The experimental units are numbered in serpentine manner

1	2	3	4	5	6
12	11	10	9	8	7
13	14	15	16	17	18
24	23	22	21	20	19
25	26	27	28	29	30

Now we select 30 (=n) distinct greater than and equal to three – digit random numbers from random number tables. In this case, number of experimental units are of two – digit, so we select three – digit random numbers from random number tables to avoid the repetition of the same number. Again, the random numbers are written and are ranked according to its increasing order. The lowest random number is denoted

by rank 1, and the highest random number is ranked by the largest number. Next, the first set of r experimental units are allocated to treatment t_1, the second set of r experimental units are allocated to treatment t_2, the third set of r experimental units are allocated to treatment t_3, and so on till all the set of r experimental units are allocated to treatment t_t. We have shown the whole procedure of selecting three – digit random numbers along with its rank in Table – 4.

Table – 4 Three – digit random numbers, its rank and treatments allotted

Three -digit random number	Rank	Treatment to be allotted
810	27	t_1
335	14	t_1
256	10	t_1
679	23	t_1
035	2	t_1
803	26	t_2
567	18	t_2
815	28	t_2
237	9	t_2
045	3	t_2
416	16	t_3
849	29	t_3
325	13	t_3
087	5	t_3
578	19	t_3
277	11	t_4
786	25	t_4
918	30	t_4
035	1	t_4
651	21	t_4
548	17	t_5
386	15	t_5
209	8	t_5
598	20	t_5
069	4	t_5
137	7	t_6
759	24	t_6
659	22	t_6
312	12	t_6
099	6	t_6

From Table – 4, we can observe that experimental unit numbers 27, 14, 10, 23, and 2 are allotted to treatment t_1, experimental unit numbers 26, 18, 28, 9, and 3 are allotted to treatment t_2, experimental unit numbers 16, 29, 13, 5, and 19 are allotted to treatment t_3, and so on. The final lay out of the completely randomized design is shown in Table – 5.

Table – 5 Completely randomized design with six treatments each is replicated five times

$1(t_4)$	$2(t_1)$	$3(t_2)$	$4(t_5)$	$5(t_3)$	$6(t_6)$
$12(t_6)$	$11(t_4)$	$10(t_1)$	$9(t_2)$	$8(t_5)$	$7(t_6)$
$13(t_3)$	$14(t_1)$	$15(t_5)$	$16(t_3)$	$17(t_5)$	$18(t_2)$
$24(t_6)$	$23(t_1)$	$22(t_6)$	$21(t_4)$	$20(t_5)$	$19(t_3)$
$25(t_4)$	$26(t_2)$	$27(t_1)$	$28(t_2)$	$29(t_3)$	$30(t_4)$

Planning of Completely randomized design

Once the observations are collected, these are planned and arranged in Table – 6.

Table – 6 Plan and arrangement of observations

Model of Completely randomized design

The linear model of the completely randomized design is defined as

Treatments				
1	2	3	…	t
Y_{11}	Y_{21}	Y_{31}	---	Y_{t1}
Y_{12}	Y_{22}	Y_{32}	…	Y_{t2}
Y_{13}	Y_{23}	Y_{33}	…	Y_{t3}
:	:	:	:	:
:	:	:	:	:
Total: t_1	t_2	t_3	…	t_t
Number of replications:r_1	r_2	r_3	…	r_t

$Y_{ij} = \mu + \tau_i + e_{ij}$ (1)

where, denotes the observations due to treatment of experimental unit, is treatment effects and is experimental error due to treatment of experimental unit.

Analysis of Variance

First of all, we test the homogeneity of the variance. If all the variances are homogenous, we proceed further for obtaining analysis of variance table, otherwise, use the transformation method to make the data homogenous.

We compute the following sum of squares for obtaining the analysis of variance table.

Treatment sum of squares = $\left[\frac{t_1^2}{r_1} + \frac{t_2^2}{r_2} + \cdots + \frac{t_t^2}{r_t}\right]$ - CF = $\sum_{i=1}^{t} \frac{t_i^2}{r_i}$ – CF,

Where, CF = $\frac{(Grand\ Total)^2}{n}$.

Total sum of squares = $(Y_{11}^2 + Y_{12}^2 + \cdots + Y_{tr}^2) - CF = \sum_{i=j=1}^{t} Y_{ij}^2$ -CF,

Error um of squares = Total sum of squares – treatment sum of squares.

These sums of squares are summarized in Table – 7.

Table – 7 Analysis of variance table of completely randomized design

Sources of variation	df	Sum of squares	Mean squares	F-ratio
Treatments	t - 1	TRSS	$TRSS/(t-1)= \sigma_t^2$	σ_t^2/σ_e^2
Error	tr - t	ESS	$ESS/(tr-1) = \sigma_e^2$	
Total	tr - 1	TSS		

Conclusion: If calculated value of F - test is greater than tabulated value of F - test at αpercent of level of significance, the test is significant. Therefore, we reject the null hypothesis. That is, there is statistically significant difference among the t treatments.

Practical example

An agricultural experimental station wishes to determine the yields of five

varieties of corn. Each type of corn is planted to four plots of land with equal fertility, and the yields, in units of 100 bushels, are shown below. In view of these results, can we conclude that the mean yields of the five varieties of corn are the same using the 5% level of significance?

Varieties

A	B	C	D	E
4	7	5	11	4
6	13	4	10	8
4	10	4	9	6
10	12	9	14	10

Here, A, B, C, D, and E are five treatments and each is replicated four times. So, t = 5, r = 4 and experimental units, n = 20.

Correction factor, CF = $\frac{(Grand\ Total)^2}{n}$. = $\frac{(160)^2}{20}$.= 1280.

Treatment sum of squares = $\left[\frac{t_1^2}{r_1} + \frac{t_2^2}{r_2} + \cdots + \frac{t_t^2}{r_t}\right]$ - CF

$= \left[\frac{(24)^2}{4} + \frac{(42)^2}{4} + \frac{(22)^2}{4} + \frac{(44)^2}{4} + \frac{(28)^2}{4}\right]$ - 1280 = 106.

Total sum of squares = $(Y_{11}^2 + Y_{12}^2 + \cdots + Y_{tr}^2) - CF$

$= [(4)^2 + (7)^2 + \cdots + (10)^2] - 1280 = 202.$

Error um of squares = Total sum of squares – treatment sum of squares

= 202 – 106 = 96

Table – 8 Analysis of variance table of completely randomized design

Sources of variation	df	Sum of squares	Mean squares	F-ratio
Treatments	4	106	106/4=26.5	26.5/6.4 = 4.14
Error	15	96	96/15 = 6.4	
Total	19	202		

Conclusion: Calculated value of F is 4.14 which is greater than tabulated value of F at five percent (0.05%) level of significance. Therefore, F- test is significant

and hence, there is statistically significant difference among the treatments.

Advantages/Disadvantages: Agricultural experiment is a field experiments so there is every possibility of data to be heterogenous. Hence, we cannot use completely randomized design in agricultural experiments. However, if we conduct the laboratory experiment then the collected observations will always be homogenous. Therefore, we always recommend to apply completely randomized design in laboratory experiments.

2. RANDOMIZED BLOCK DESIGNS

Randomized block designs are used in greater extent in agricultural experiments because of low cost, least number of treatments, least number of replications and block sizes as well as simple and easy in analysing the data.

Let there are t treatments, each treatment is replicated r times. So, total number of experimental units is tr. Which is arranged in b blocks such that each block contains k treatments.

A randomized block design involves subjects being split into two groups (or blocks) such that the variation within the groups (according to the chosen matching variables) is less than the variation between the groups.

The block randomization method is designed to randomize subjects into groups that result in equal sample sizes. This method is used to ensure a balance in sample size across groups over time.

Generalized randomized block design

In case of randomized experiments, generalized randomized block designs are used to study the interaction effects between blocks and treatments. In case of generalized randomized block designs (GBRD), each treatment is replicated at least two times in each block. These replications allow us to estimate the interaction effects and then testing that interaction effect in the linear model.

Planning and arrangement of data in randomized block design

Let us collect the tr observations and then arrange them in two-way table

according to blocks(replications) and treatments. Arrangement of tr data in two-way table is shown in Table – 9.

Table – 9 Plan and arrangement of observations for randomized block design

Treatment	Replication					
	1	2	3	..	r	Total
1	Y_{11}	Y_{12}	Y_{13}	..	Y_{1r}	T_1
2	Y_{21}	Y_{22}	Y_{23}	..	Y_{2r}	T_2
3	Y_{31}	Y_{32}	Y_{33}	..	Y_{3r}	T_3
:	..	..	..	..	..	..
t	Y_{t1}	Y_{t2}	Y_{t3}	::	Y_{tr}	T_t
Total	R_1	R_2	R_3	..	R_t	

Model of randomized block design

For every experimental design, there is a mathematical model that accounts for all of the independent and extraneous variables that affect the dependent variable. The mathematical model for randomized block design follows:

$$Y_{ij} = \mu + \tau_i + \beta_j + e_{ij} \qquad (2)$$

where, is the dependent variable denotes the observations due to treatment received in block, is treatment effects, β_j is the effect of j^{th} block and e_{ij} is experimental error (i.e., the effect of all other extraneous variables) due to i^{th} treatment received in j^{th} block.

For this model, it is assumed that e_{ij} is normally and independently distributed with a mean of zero and a variance of σ_e^2. The mean (μ) is constant.

Statistical hypothesis

With a randomized block experiment, it is possible to test both block (β_j) and treatment (τ_i) effects. Here the null hypothesis (H_0) and alternate hypothesis (H_1) for each effect follows:

H_0: $\beta_j = 0$ for all j

H_1: $\beta_j \neq 0$ for some j

H_0: $\tau_i = 0$ for all i

H_1: $\tau_i \neq 0$ for some i

With a randomized block design, the main hypothesis test of interest is the test of the treatment effect(s). Block effects are of less intrinsic interest, because a blocking variable is thought to be a nuisance variable that is only included in the experiment to control for a potential source of undesired variation.

Analysis of Variance

First of all, we test the homogeneity of the variance. If all the variances are homogenous, we proceed further for obtaining analysis of variance table, otherwise, use the transformation method to make the data homogenous.

We compute the following sum of squares for obtaining the analysis of variance table.

Treatment sum of squares = $\left[\frac{T_1^2}{r} + \frac{T_2^2}{r} + \cdots + \frac{T_t^2}{r}\right]$ - CF = $\frac{1}{r}\sum_{i=1}^{t} T_i^2$ – CF,

Where, CF = $\frac{(Grand\ Total)^2}{n}$.

Total sum of squares = $(Y_{11}^2 + Y_{12}^2 + \cdots + Y_{tr}^2) - CF = \sum_{i=j=1}^{t} Y_{ij}^2$ - CF,

Block sum of squares = $\frac{1}{t}\sum_{j=1}^{b} B_j^2$ – CF

Error um of squares = Total sum of squares – treatment sum of squares – Block sum of squares.

These sums of squares are summarized in Table – 10.

Table – 10 Analysis of variance table of completely randomized design

Sources of variation	df	Sum of squares	Mean squares	F-ratio
Treatment	t - 1	TRSS	$TRSS/(t-1) = \sigma_t^2$	σ_t^2/σ_e^2
Block	b - 1	BSS	$BSS/(b-1) = \sigma_b^2$	σ_b^2/σ_e^2
Error	(t – 1)(b – 1)	ESS	$ESS/[(t-1)(b-1)] = \sigma_e^2$	
Total	tr - 1	TSS		

Conclusion: If calculated value of F - test is greater than tabulated value of F - test at α percent of level of significance, the test is significant. Therefore, we reject the null hypothesis. That is, there is statistically significant difference among the t treatments.

P-value

In an experiment, a P-value is the probability of obtaining a result more extreme than the observed experimental outcome, assuming the null hypothesis is true.

With analysis of variance for a randomized block experiment, the *F* ratios are the observed experimental outcomes that we are interested in. So, the P-value would be the probability that an *F*- ratio would be more extreme (i.e., bigger) than the actual *F* - ratio computed from experimental data.

Example 2: Three different washing solutions are being compared to study their effectiveness in retarding bacteria growth in 5-gallon milk containers. The analysis is done in a laboratory, and only three trials can be run on any day. Because days could represent a potential source of variability, the experimenter decides to use a randomized block design. Observations are taken for four days, and the data are shown here. Analyse the data from this experiment and draw conclusions. In this example, the blocking factor is the day. The treatment is "solution". We have three types of solutions and four levels for "day." The data is shown in Table – 11.

Table – 11 Randomized block design with three treatments each repeated in four blocks

Solutions	**Days**			
	1	2	3	4
1	13	22	18	39
2	16	24	17	44
3	5	4	1	22

Here, solutions 1, 2 and 3 are three treatments and each is repeated in four

blocks. So, t = 3, b = 4 and experimental units, n = 12.

Correction factor, CF = $\frac{(Grand\ Total)^2}{n}$. = $\frac{(225)^2}{12}$.= 4218.75.

Treatment sum of squares = $\left[\frac{t_1^2}{r_1} + \frac{t_2^2}{r_2} + \cdots + \frac{t_t^2}{r_t}\right]$ - CF

= $\left[\frac{(92)^2}{4} + \frac{(101)^2}{4} + \frac{(32)^2}{4}\right]$ - 4218.75 = 703.5.

Block sum of squares = $\frac{1}{t}\sum_{j=1}^{b} B_j^2$ – CF

= $\left[\frac{(34)^2}{3} + \frac{(50)^2}{3} + \frac{(36)^2}{3} + \frac{(105)^2}{3}\right]$ - 4218.75 = 1106.92

Total sum of squares = $(Y_{11}^2 + Y_{12}^2 + \cdots + Y_{tr}^2) - CF$

= $[(13)^2 + (22)^2 + \cdots + (22)^2]$ – 4218.75 = 1862.25.

Error um of squares = Total sum of squares – treatment sum of squares – Block sum of squares

= 1862.25 - 703.5 - 1106.92 = 51.83.

Analysis of variance table is shown in Table – 12.

Table – 12 Analysis of variance table od randomized block design

SV	DF	SS	MS	F-Ratio	p-value
Treatments	2	703.5	351.75	40.76***	0.0003
Blocks	3	1106.92	368.97	42.75	
Error	6	51.83	8.63		
Total	11	1862.25			

Conclusion: Calculated value of F is 40.76 for treatment variation which is greater than tabulated value of F at 0.1 percent (0.001%) level of significance. Therefore, F- test is significant. Again p-value of treatment variation is 0.0003 which is less than 0.05, so treatment is significant. Hence, there is statistically significant difference among the treatments

Advantages and disadvantages of randomized block designs

Following are the advantages and disadvantages of the randomized block designs:

1. Randomized block design increases the precision of the experiment.

2. Randomized block design is more efficient than complete randomized design.
3. Any number of blocks can be considered for the randomized block designs.
4. The analysis of variance is easy and simple, even though some observations are missing, which can be estimated using missing plot techniques.

Disadvantages: Let the number of treatments is increased, consequently block size will increase and hence, it is not possible to maintain the homogeneity of the data. So, design will be less precise as experimental design will be large.

3. Latin Square Designs

The Latin square design applies when there are repeated exposures/treatments and two other factors. This design avoids the excessive numbers required for full three-way ANOVA.

An example of a Latin square design is the response of 5 different fertilizers (factor 1) to 5 different treatments (repeated blocks A to E) applied on 5 different variety of seeds (factor 2):

		Fertilizers				
		1	2	3	4	5
	1	A	E	C	D	B
Seeds	2	E	B	A	C	D
	3	C	D	E	B	A
	4	D	C	B	A	E
	5	B	A	D	E	C

This special sort of balancing means that the systematic variation between rows, or similarity between columns, does not affect the comparison of treatments.

The following example taken from Mead et al. (2003) illustrates the concept of Latin squares design.

An experiment to investigate the effects of various dietary starch levels on milk production was conducted on four cows. The four diets, T_1, T_2, T_3, and T_4, (in order

of increasing starch equivalent), were fed for three weeks to each cow and the total yield of milk in the third week of each period was recorded (i.e. third week to minimize carry-over effects due to the use of treatments administered in a previous period). That is, the trial lasted 12 weeks since each cow received each treatment, and each treatment required three weeks. The investigator felt strongly that time period effects might be important (i.e earlier periods in the experiment might influence milk yields differently compared to later periods). Hence, the investigator wanted to block on both cow and period. However, each cow cannot possibly receive more than one treatment during the same time period; that is, all possible cow-period blocking combinations could not logically be considered.

The Latin square is probably under used in most fields of research because text book examples tend to be restricted to agriculture, the area which spawned most original work on ANOVA. Agricultural examples often reflect geographical designs where rows and columns are literally two dimensions of a grid in a field. Rows and columns can be any two sources of variation in an experiment. In this sense a Latin square is a generalisation of a randomized block design with two different blocking systems.

Planning and arrangement of data in Latin squares design

Suppose we have 4 treatments (namely: A, B, C, and D), then it means that we have Number of Treatments = Number of Rows = Number of Columns =4.

Arrangement of 16 data in two-way table for Latin squares design is shown in Table – 13.

Table – 13 Plan and arrangement of observations for Latin squares design

Rows	Columns			
1	A	B	C	D
	Y_{111}	Y_{122}	Y_{133}	Y_{144}
2	B	C	D	A
	Y_{212}	Y_{223}	Y_{234}	Y_{241}
3	C	D	A	B
	Y_{313}	Y_{324}	Y_{331}	Y_{342}
4	D	A	B	C
	Y_{414}	Y_{421}	Y_{432}	Y_{443}

The mathematical model for Latin squares design follows:

$$Y_{ijk} = \mu + r_i + c_j + t_k + e_{ijk} \quad (3)$$

where, , is the observation for the experimental unit in the row, column and the treatment effect, is the row effects, is the column effects, is the treatment effects and is experimental error, i = 1,2, .. , t; j = 1, 2, … , t; and k = 1, 2, …, t, as number of rows = number of columns = number of treatments = t for Latin squares design.

For this model, it is assumed that is normally and independently distributed with a mean of zero and a variance of . The mean (μ) is constant.

Statistical hypothesis

With a Latin squares design, it is possible to test both row (r_i), column (c_j) and treatment (τ_k) effects. Here the null hypothesis (H_0) and alternate hypothesis (H_1) for each effect follows:

H_0: $r_i = 0$ for all i

H_1: $r_i \neq 0$ for some i

H_0: $c_j = 0$ for all j

H_1: $c_j \neq 0$ for some j

H_0: $\tau_k = 0$ for all k

H_1: $\tau_k \neq 0$ for some k

With a Latin squares design, the main hypothesis test of interest is the test of the treatment effect(s); row effects, column effects are of less intrinsic interest, because two blocking variables are thought to be a nuisance variable that is only included in the experiment to control for a potential source of undesired variation.

Analysis of Variance

First of all, we test the homogeneity of the variance. If all the variances are homogenous, we proceed further for obtaining analysis of variance table, otherwise, use the transformation method to make the data homogenous.

We compute the following sum of squares for obtaining the analysis of variance table.

Treatment sum of squares = $\left[\frac{T_1^2}{t} + \frac{T_2^2}{t} + \cdots + \frac{T_t^2}{t}\right]$ - CF = $\frac{1}{t}\sum_{i=1}^{t} T_i^2$ – CF,

Where, CF = $\frac{(Grand\ Total)^2}{t^2}$.

Total sum of square = $\sum_{i=j=k=1}^{t} Y_{ijk}^2$ - CF,

Row sum of squares = $\frac{1}{t}\sum_{i=1}^{t} R_i^2$ – CF

Column sum of squares = $\frac{1}{t}\sum_{j=1}^{t} C_j^2$ – CF

Error um of squares = Total sum of squares – treatment sum of squares – Row sum of squares – Column sum of squares.

These sums of squares are summarized in Table – 14.

Table – 14 Analysis of variance table of Latin squares design

Sources of variation	df	Sum of squares	Mean squares	F-ratio
Treatment	t - 1	TRSS	$TRSS/(t-1) = \sigma_t^2$	σ_t^2/σ_e^2
Row	t - 1	RSS	$RSS/(t-1) = \sigma_R^2$	σ_R^2/σ_e^2
Column	t - 1	CSS	$CSS/(t-1) = \sigma_C^2$	σ_C^2/σ_e^2
Error	(t – 1)(t – 2)	ESS	$ESS/[(t-1)(t-2)] = \sigma_e^2$	
Total	t^2 - 1	TSS		

Conclusion: If calculated value of F - test is greater than tabulated value of F - test at α percent of level of significance, the test is significant. Therefore, we reject the null hypothesis. That is, there is statistically significant difference among the t treatments.

Example – 3: Armitage quotes a paper which reported an experiment that had been designed as a Latin square. The skins of rabbits' backs were inoculated with a diffusing factor in six separate sites. Six rabbits were therefore used and the order in which the sites were inoculated was done six different ways. The outcome measured was area of blister (cm²). The overall objective was to see whether or not the order

of administration affected this outcome. The experimental design and data are represented in the Latin square shown in Table -15.

Table – 15 Latin squares design with six rows, six columns, and six treatments

Position	Rabbit					
	1	2	3	4	5	6
a	III (7.9)	V (8.7)	IV (7.4)	I (7.4)	VI (7.1)	II (8.2)
b	IV (6.1)	II (8.2)	VI (7.7)	V (7.1)	III (8.1)	I (5.9)
c	I (7.5)	III (8.1)	V (6.0)	VI (6.4)	II (6.2)	IV (7.5)
d	VI (6.9)	I (8.5)	III (6.8)	II (7.7)	IV (8.5)	V (8.5)
e	II (6.7)	IV (9.9)	I (7.3)	III (6.4)	V (6.4)	VI (7.3)
f	V (7.3)	VI (8.3)	II (7.3)	IV (5.8)	I (6.4)	III (7.7)

Correction factor, CF = $\frac{(Grand\ Total)^2}{t^2}$. = $\frac{(265.2)^2}{36}$.= 1953.64.

Treatment sum of squares = $\left[\frac{t_1^2}{r_1} + \frac{t_2^2}{r_2} + \cdots + \frac{t_t^2}{r_t}\right]$ - CF

= $\left[\frac{(43)^2}{6} + \frac{(44.3)^2}{6} + \frac{(45)^2}{6} + \frac{(45.2)^2}{6} + \frac{(44)^2}{6} + \frac{(43.7)^2}{6}\right]$ - 1953.64

= 0.563333.

Row sum of squares = $\frac{1}{t}\sum_{i=1}^{t} R_i^2$ – CF

= $\left[\frac{(46.7)^2}{6} + \frac{(43.1)^2}{6} + \frac{(41.7)^2}{6} + \frac{(46.9)^2}{6} + \frac{(44)^2}{6} + \frac{(42.8)^2}{6}\right]$ - 1953.64

= 3.833333.

Column sum of squares = $\frac{1}{t}\sum_{j=1}^{t} C_j^2$ – CF

$\left[\frac{(42.4)^2}{6} + \frac{(51.7)^2}{6} + \frac{(42.5)^2}{6} + \frac{(40.8)^2}{6} + \frac{(42.7)^2}{6} + \frac{(45.1)^2}{6}\right]$ - 1953.64

= 12.83333.

Total sum of squares = $(Y_{111}^2 + Y_{122}^2 + \cdots + Y_{666}^2) - CF$

= $[(7.9)^2 + (8.7)^2 + \cdots + (7.7)^2]$ – 1953.64 = 30.36.

Error um of squares = Total sum of squares – treatment sum of squares – Row

sum of squares – Column sum of squares = 30.36 - 0.563333 - 3.833333 - 12.83333 = 13.13.

Analysis of variance table is shown in Table – 16.

Table – 16 Analysis of variance table of Latin squares design

SV	DF	SS	MS	F-Ratio	p-value
Treatments	5	0.563333	0.11266	0.171617	0.97013^{NS}
Row	5	3.833333	0.76666	1.167809	0.359187
Column	5	12.83333	2.56666	3.909622	0.012352
Error	20	13.13	0.6565		
Total	35	30.36			

Conclusion: Calculated value of F is 0.171617 for treatment which is less than tabulated value of F at five percent (0.05%) level of significance. Therefore, F- test is insignificant. Again p-value of treatment variation is 0.97013 which is greater than 0.05, so treatment is insignificant. Hence, the order of administration do not affected this outcome

Advantages of the Latin squares Design

1. You can control variation in two directions.
2. Hopefully you increase efficiency as compared to the CRD and RBD.

Disadvantages of the Latin squares Design

1. The number of treatments must equal the number of replicates.
2. The experimental error is likely to increase with the size of the square.
3. Small squares have very few degrees of freedom for experimental error.
4. You can't evaluate interactions between: (a) Rows and columns, (b) Rows and treatments, (c) Columns and treatments.

Note that a Latin Square is an incomplete design, which means that it does not include observations for all possible combinations of i, j and k.

Mead, R., R.N. Curnow, and A.M. Hasted. 2003. Statistical methods in agriculture and experimental biology. Chapman and Hall/CRC, Boca Raton, FL.

4. Factorial Experiment

If the researchers are interested to study the effects of all possible combinations of all the levels of all factors or some of the factors, then the basic designs does not serve purpose. So, we another type of experimental designs. This is called as factorial designs. Factorial design provides the response of main effects and interaction effects instead of treatment effects.

In this section, we discuss some methods for estimating main effects and interaction effects for 2^n factorial experiments.

4.1 Ghosh method: This method is based on Hadamard matrix. So first of all, we define Hadamard matrix.

Definition: **Hadamard Matrix**. A matrix H_n is said to be Hadamard matrix, if $H_n H_n^{'} = H_n' H_n = n.I_n$ with $H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$; where n is divisible by 4.

Similarly, $H_4 = \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

Following are the steps for estimating the main and interaction effects due to Ghosh method. Construct a 2^n (n = 2, 3, 4, …) Hadamard Matrix. Delete first row of this matrix. Write treatment combinations $a_0b_0, a_1b_0, a_0b_1, and\ a_1b_1$ from last column to first column. Add one more column. Write main effects and interaction effects under added first column. Finally, find the estimate of Main/Interaction effect.

Example - 4 Consider a 2^2 factorial experiments with two factors, say, **A** and B each are at two levels. Here we have to prepare a Hadamard matrix with 4 rows and 4 columns. Which are given below.

$$H_4 = \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

Delete element of first row and in that place write the treatment combinations as following

Table – 17 Main effects and interaction effects

	a_1b_1	a_0b_1	a_1b_0	a_0b_0
A	1	-1	1	-1
B	1	1	-1	-1
AB	1	-1	-1	1

Main effect of A = $a_1b_1 - a_0b_1 + a_1b_0 - a_0b_0$,

Main effect of B = $a_1b_1 + a_0b_1 - a_1b_0 - a_0b_0$, and

Interaction effect of AB = $a_1b_1 - a_0b_1 - a_1b_0 + a_0b_0$

Example - 5 Consider 2^3 factorial experiment with three factors A, B and C each at two levels. Here we have to prepare a Hadamard matrix with 8 rows and 8 columns. Which are given below.

$$H_8 = \begin{bmatrix} H_4 & H_4 \\ H_4 & -H_4 \end{bmatrix} \quad \text{or,} \quad H_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix}$$

Delete element of first row and in that place write the 8 treatment combinations from reverse order as shown in Table 18.

Table – 18 Showing the sign of eight treatment combinations for respective main/interaction effects

	$a_1b_1c_1$	$a_0b_1c_1$	$a_1b_0c_1$	$a_0b_0c_1$	$a_1b_1c_0$	$a_0b_1c_0$	$a_1b_0c_0$	$a_0b_0c_0$
A	1	-1	1	-1	1	-1	1	-1
B	1	1	-1	-1	1	1	-1	-1
AB	1	-1	-1	1	1	-1	-1	1
C	1	1	1	1	-1	-1	-1	-1
AC	1	-1	1	-1	-1	1	-1	1
BC	1	1	-1	-1	-1	-1	1	1
ABC	1	-1	-1	1	-1	1	1	-1

4.2 General Analysis of 2^k Factorial experiment in r replications

Let us consider k factors each at two (low and high) levels. However, the entire experiment is repeated r times. Degrees of freedom, Sum of squares and Mean square for the respective main effects and interaction effects for 2^k factorial experiment is shown in Table – 18.

Table – 18 Analysis of variance of factorial experiments

Sources of Variation	Degrees of Freedom	Sum of squares	Mean Squares	F-ratio
Replication	r - 1	S_R^2	$S_R^2/(r-1) = \sigma_R^2$	σ_R^2/σ_e^2
	K Main effects			
A	1	S_A^2	$S_A^2/1 = \sigma_A^2$	σ_A^2/σ_e^2
B	1	S_B^2	$S_B^2/1 = \sigma_B^2$	σ_B^2/σ_e^2
:	:	:	;	:
K	1	S_K^2	$S_K^2/1 = \sigma_K^2$	σ_K^2/σ_e^2
	Two factor Interactions			
AB	1	S_{AB}^2	$S_{AB}^2/1 = \sigma_{AB}^2$	σ_{AB}^2/σ_e^2
AC	1	S_{AC}^2	$S_{AC}^2/1 = \sigma_{AC}^2$	σ_{AC}^2/σ_e^2
:	:	:	;	:
JK	1	S_{JK}^2	$S_{JK}^2/1 = \sigma_{JK}^2$	σ_{JK}^2/σ_e^2
	Three Factor Interactions			
ABC	1	S_{ABC}^2	$S_{ABC}^2/1 = \sigma_{ABC}^2$	$\sigma_{ABC}^2/\sigma_e^2$
ABD	1	S_{ABD}^2	$S_{ABD}^2/1 = \sigma_{ABD}^2$	$\sigma_{ABD}^2/\sigma_e^2$
:	:	:	;	:
IJK	1	S_{IJK}^2	$S_{IJK}^2/1 = \sigma_{IJK}^2$	$\sigma_{IJK}^2/\sigma_e^2$
K^{th} Factor interaction ABC….JK	1	$S_{ABC..K}^2$	$S_{ABC...K}^2/1 = \sigma_{ABC..K}^2$	$\sigma_{ABC...K}^2/\sigma_e^2$
Error	$(2^k - 1)(r - 1)$	S_e^2	$S_e^2/(2^k - 1)(r - 1) = \sigma_e^2$	
Total	$r2^k - 1$			

Advantages of factorial experiments

Many people examine the effect of only a single factor or variable. However, compared to such one-factor-at-a-time (OFAT) experiments, factorial experiments have following advantages

1. Factorial designs are more efficient than OFAT experiments. They provide more information at similar or lower cost. They can find optimal conditions faster than OFAT experiments.

2. Factorial designs allow additional factors to be examined at no additional cost.

3. When the effect of one factor is different for different levels of another factor, it cannot be detected by an OFAT experiment design. Factorial designs are required to detect such interaction. When interactions are present then use of OFAT can lead to serious misunderstanding of how the response changes with the factors.

4. Factorial designs allow the effects of a factor to be estimated at several levels of the other factors, yielding conclusions that are valid over a range of experimental conditions.

4.3 General Analysis of 3^k Factorial experiment in r replications

Let us consider k factors each at three levels. However, the entire experiment is repeated r times. Degrees of freedom, Sum of squares and Mean squares for the respective main effects and interaction effects for 3^k factorial experiment are shown in Table – 19.

Table – 19 Showing the ANOVA Table of factorial experiment**.**

Sources of Variation	Degrees of Freedom	Sum of squares	Mean Squares	F-ratio
Replication	r - 1	S_R^2	$S_R^2/(r-1) = \sigma_R^2$	σ_R^2/σ_e^2
K Main effects				
A	2	S_A^2	$S_A^2/2 = \sigma_A^2$	σ_A^2/σ_e^2
B	2	S_B^2	$S_B^2/2 = \sigma_B^2$	σ_B^2/σ_e^2
:	:	:	;	:
K	2	S_K^2	$S_K^2/2 = \sigma_K^2$	σ_K^2/σ_e^2
Two factor Interactions				
AB	4	S_{AB}^2	$S_{AB}^2/4 = \sigma_{AB}^2$	σ_{AB}^2/σ_e^2
AC	4	S_{AC}^2	$S_{AC}^2/4 = \sigma_{AC}^2$	σ_{AC}^2/σ_e^2
:	:	:	;	:
JK	4	S_{JK}^2	$S_{JK}^2/4 = \sigma_{JK}^2$	σ_{JK}^2/σ_e^2
Three Factor Interactions				
ABC	8	S_{ABC}^2	$S_{ABC}^2/8 = \sigma_{ABC}^2$	$\sigma_{ABC}^2/\sigma_e^2$
ABD	8	S_{ABD}^2	$S_{ABD}^2/8 = \sigma_{ABD}^2$	$\sigma_{ABD}^2/\sigma_e^2$
:	:	:	;	:
IJK	8	S_{IJK}^2	$S_{IJK}^2/8 = \sigma_{IJK}^2$	$\sigma_{IJK}^2/\sigma_e^2$
$(k-1)^{th}$ Factor Interaction	:	:	$S_{ABC\ldots K}^2/8 = \sigma_{ABC\ldots K}^2$	$\sigma_{ABC\ldots K}^2/\sigma_e^2$
k^{th}Factor interaction ABC….JK	2^k	$S_{ABC..K}^2$	$S_{ABC}^2/1 = \sigma_{ABC}^2$	$\sigma_{ABC}^2/\sigma_e^2$
Error	$(3^k - 1)(r-1)$	S_e^2	$S_e^2/(3^k-1)(r-1) = \sigma_e^2$	
Total	$r3^k - 1$			

4.4 General Analysis of S^k Factorial experiment in r replications

Let us consider k factors each at three levels. However, the entire experiment is repeated r times. Degrees of freedom, Sum of squares and Mean squares for the respective main effects and interaction effects for S^k factorial experiment are shown in Table – 20.

Table – 20 Showing the sum of squares of factorial experiment

Sources of Variation	Degrees of Freedom	Sum of squares	Mean Squares	F-ratio
Replication	r - 1	S_R^2	$S_R^2/(r-1)=\sigma_R^2$	σ_R^2/σ_e^2
		K Main effects		
A	$(S-1)$	S_A^2	$S_A^2/(S-1)=\sigma_A^2$	σ_A^2/σ_e^2
B	$(S-1)$	S_B^2	$S_B^2/(S-1)=\sigma_B^2$	σ_B^2/σ_e^2
:	:	:	;	:
K	$(S-1)$	S_K^2	$S_K^2/(S-1)=\sigma_K^2$	σ_K^2/σ_e^2
		Two factor Interactions		
AB	$(S-1)(S-1)$	S_{AB}^2	$S_{AB}^2/(S-1)*S-1)=\sigma_{AB}^2$	σ_{AB}^2/σ_e^2
AC	$(S-1)(S-1)$	S_{AC}^2	$S_{AC}^2/(S-1)*S-1)=\sigma_{AC}^2$	σ_{AC}^2/σ_e^2
:	:	:	;	:
JK	$(S-1)(S-1)$	S_{JK}^2	$S_{JK}^2/(S-1)*S-1)=\sigma_{JK}^2$	σ_{JK}^2/σ_e^2
		Three Factor Interactions		
ABC	(S–1)(S–1)(S–1)	S_{ABC}^2	$S_{ABC}^2/(S-1)(S-1)(S-1)=\sigma_{ABC}^2$	$\sigma_{ABC}^2/\sigma_e^2$
ABD	(S–1)(S–1)(S–1)	S_{ABD}^2	$S_{ABD}^2/(S-1)(S-1)(S-1)=\sigma$	$\sigma_{ABD}^2/\sigma_e^2$
:	:	:	;	:
IJK	(S–1)(S–1)(S–1)	S_{IJK}^2	$S_{IJK}^2/(S-1)(S-1)(S-1)=\sigma_{IJK}^2$	$\sigma_{IJK}^2/\sigma_e^2$
$(K-1)^{th}$ Factor Interaction	:	:	:	:
K^{th} Factor interaction ABC….JK	$(S-1)^k$	$S_{ABC..K}^2$	$S_{ABC..K}^2/(s-1)^k=\sigma_{ABC..K}^2$	$\sigma_{ABC..K}^2/\sigma$
Error	$(S^k-1)(r-1)$	S_e^2	$S_e^2/(S^k-1)(r-1)=\sigma_e^2$	
Total	rS^k-1			

5. Confounding in Factorial experiments

When number of factors as well as level of factors are large, experimental will be very large. In this situation, it is not possible to maintain the homogeneity of data. So, factorial designs fail to provide the right decision. Hence, we need another experimental design, which we called confounding in factorial design. This design reduces the block sizes so as to have homogeneity in the blocks.

Here, we discuss the analysis of variance table only for 2^3 factorial experiments confounded into two blocks per replication with block sizes four.

5.1 2^3 confounded factorial experiments into 2^2 block size with r replications

Table – 21 Showing the analysis of variance table of 2^2 confounded factorial experiments into 2^2 block size with r replications

Sources of Variations	**DF**	**SS**	**MS**	**F-Ratio**
Replication	r-1	$\sum R_i^2/2^3 - CF = S_r^2$	$S_r^2/(r-1) = \sigma_r^2$	σ_r^2/σ_e^2
NPK	1	$[NPK]^2/8r - CF = S_{NPK}^2$	$S_{NPK}^2/1 = \sigma_{NPK}^2$	$\sigma_{NPK}^2/\sigma_e^2$
Treatment	6	$\sum t_i^2/r - CF = S_t^2$	$S_t^2/6 = \sigma_t^2$	σ_t^2/σ_e^2
Main effect				
N	1	$[N]^2/8r = S_N^2$	$S_N^2/1 = \sigma_N^2$	σ_N^2/σ_e^2
P	1	$[P]^2/8r = S_P^2$	$S_P^2/1 = \sigma_P^2$	σ_P^2/σ_e^2
K	1	$[K]^2/8r = S_K^2$	$S_K^2/1 = \sigma_K^2$	σ_K^2/σ_e^2
Interaction effect				
NP	1	$[NP]^2/8r = S_{NP}^2$	$S_{NP}^2/1 = \sigma_{NP}^2$	σ_{NP}^2/σ_e^2
NK	1	$[NK]^2/8r = S_{NK}^2$	$S_{NK}^2/1 = \sigma_{NK}^2$	σ_{NK}^2/σ_e^2
PK	1	$[PK]^2/8r = S_{PK}^2$	$S_{PK}^2/1 = \sigma_{PK}^2$	σ_{PK}^2/σ_e^2
Error	7r - 6	By subtraction $= S_e^2$	$S_e^2/(7r-6) = \sigma_e^2$	
Total	8r-1	$\sum\sum Y_{ij}^2 - CF$		

Where CF = $\frac{(Grand\ total)^2}{2^3 r}$

REMARKS 3.1

(a) In case of confounding experiment, the Main effect or interaction effect which are confounded, cannot be estimated. That is, we are losing total information on that confounded main effects or interaction effects.

(b) In a confounding factorial experiment, there will be two or more than two blocks per replication while in factorial experiment there is one and only one block per replication.

(c) We will always try to confound higher - order interaction because higher - order interaction has minimum contribution, i.e., we are losing less information as

they are not important. For example, we have factorial experiments with five factors. Now consider a higher - order interaction, say, ABCDE (which is a five-factor interaction). In this experiment, contribution of main effects and two - factor interaction effects are more important compare to three, four and five - factor interactions and hence, we do not need to lose any information on main effects and two - factors interactions. Therefore, we will confound the higher - order interaction ABCDE only.

1. Confounding in factorial experiments with single replication.

So far, we discussed the confounding factorial experiments with two and more than two replications. We always prefer to conduct the experiments with two or more than two replications. Because this system provides more precise and efficient estimates. However, there are many circumstances where resources are not available for more than one replication. In such situations we are bound to analysis the data with single replication. The drawback of the single replication is that we cannot derive the estimate of error variance. In many experiments with single replication, it may happen that higher - order interactions are negligible. Therefore, sum of squares of such higher - order interactions should be considered as the sum of the squares of error variations. We discuss it for both 2^n and 3^n factorial experiments through some examples.

6.1 Confounding in 2^n factorial experiment with single replication.

Example – 6 We consider 2^5 confounding factorial experiments in to a block of size 8 by confounding two independent interactions ABD and CDE with block. Its generalized interaction is ABCE. This experiment provides four blocks each of size eight per replication. Since total degree of freedom is 31, so error degree of freedom will be negligible. Therefore, it is advisable to pull sum of squares of higher order interaction with error sum of squares. This is summarized in Table – 22.

Table – 22 ANOVA table with single replication

Sources of Variations	Degree of Freedom
Block	3
Main effects	5
Two - factor Interactions	10
Error (Comprises of all three - factor Interactions, ABCE and any two other higher - order interactions)	13
Total	31

6.2 Confounding in 3^n factorial experiment with single replication.

Example – 7 We consider 3^3 confounding factorial experiments in to a block of size 9 by confounding one independent interactions ABC with block. This experiment provides three blocks each of size nine per replication. Since total degree of freedom is 26, so error degree of freedom will be negligible. Therefore, it is advisable to pull sum of squares of higher order interaction with error sum of squares. This is summarized in Table – 23.

Table – 23 ANOVA table with single replication

Sources of Variations	Degree of Freedom
Block	2
Main effects	6
Two factor Interactions	12
Error (Comprises of three - factor Interaction contrast, leaving any two contrasts of three - factor interactions, except ABC and $A^2B^2C^2$ *as these are confounded*)	6
Total	26

ACKNOWLEDGEMENTS

I thank the referee for reviwing this paper.

REFERENCES

[1] **Berry (1994)** Latin Square Design, Field trials in which the experimental error has two fertility gradients running, page no 236, Google search.

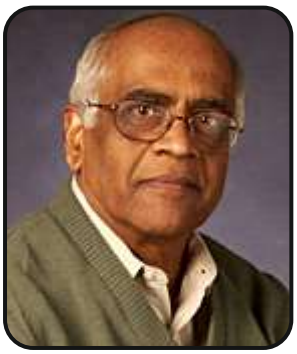
[2] **Mead, R., R.N. Curnow, and A.M. Hasted. 2003.** Statistical methods in agriculture and experimental biology. Chapman and Hall/CRC, Boca Raton, FL

[3] **R. Rangaswamy (1995).** A text book of Agricultural Statistics, New age international publication limited (Wiley eastern limited), New Delhi.

BIOGRAPHY

PROF. ANANT KSHIRSAGAR*

H. D. BUDHBHATTI**



Professor Anant Madhav Kshirsagar was born on August 19, 1931 in Sangali (Maharashtra). He received his matriculation at an early age of 12 years. (This was due to the private school where ONE can be promoted earlier keeping his excellent academic records.

He joined Fergusson College, Pune for his study of B.Sc. with Maths as Principal subject and Physics as subsidiary subject. He passed B.Sc. exam in 1948 at the age of 16, securing 83% marks and topping the list. He received Pragji Thakarsi Mulasi scholarship and also R. P. Paranjape prize. At that time under the guidance of Prof. P. V. Sukhatme, University of Bombay was offering Masters Degree course in Statistics. This was initiated by Prof. M. C. Chakrabarty who had built up this department. Kshirsagar joined B. U. and received his M.Sc. degree in statistics in 1951.

He started his academic career by joining B. U. after his post graduate degree. He had started as Demonstrator and then became lecturer. He was teaching Lincer Models, Statistical Inference, Industrial Statistics, Design of experiments, sample surveys and Multivariate analysis.

Kshirsagar was basically more interested in Multivariate analysis. He was advised by Prof. P. V. Sukhatme to approach Prof. Bartlett who was in Manchester. He was

* Adapted from Wikepedia (the free encyclopedia) and other related sources. We express our sincere thanks for their assistance.

** Ex. CSO, GSRTC, Ahmedabad (email : hdbudhbhatti@gmail.com)

I thank the referee for reviewing this article.
(rcd. March '22 / rvd. June '22)

registered for Ph.D. degree in Multivariate analysis under the guidance of Prof. Bartlett.

He received teaching jobs at two places - first at Manchester and then at the University College, London in 1960. He got tutor's job at the university. He was fortunate to meet distinguished statisticians such as **E. S. Pearson, Barton, F. N. David, D.R.Cox, G.Bernard** etc. in London. He published number of papers in topics of design of experiments, multivariate analysis, Markov renewal process etc.

He received Ph.D. degree in 1961 from Manchester University. Later on, he also received D.Sc. degree in statistics from Manchester University in 1976. His works in the areas of connonical correlations were much appreciated.

After completing Ph.D. he came back to Bombay University and served from 1961 to 1963. He was invited by UPSC for the post of senior scientific officer - Grade I in the Defence Science Laboratory (DSL), Delhi. He served there for 5 years. In 1968, he received good placement in southern Methodist University and served there upto 1972.

His works were much appreciated by **Prof. J.N.K.Rao, Dr. D.B.Own, Dr. R.L.Mason, Dr. W.R. Sahucany** etc.

Dr. Own made him associate editor of the **Journal Communications in Statistics.**

In 1972 he joined Texas A and M University in 1977, he joined Biostatistics department of the University of Michigan.

Dr. Owen encouraged him to write a textbook on Multivariate Analysis under the publisher Marcel Dekkar, which came out successfully very soon.

He also wrote a book on Linear models and Growth Curves (Jointly with Dr. W.B. Surith).

prof. Kshirsagar had guided about 35 students for doctorate research work. These students belong to countries like US, Taiwan, Korea, Chma, Portugal, Phillipines, Argentina, Sudan, Australia and India.

Most of them have worked in multivariate analysis, designs of experiements Discriminant analysis etc.

He retired from his job in 2015 after working from a long span period of 50 years. He has one desire to revise and enlarge his mentor Prof. Chakrabarty's book **- Mathematics of Experiemental Designs.**

He is a very polite and decent academician with tremendous energy source for academic working in his choice areas and he is always ready to encourage all young statisticians and academicians for advancement of their works. He spends some time in India and some time in Anna Arbor.

We all can get inspiration and encouragement from this deignitory in our statistics field.

RESEARCH INTERESTS & PROJECTS

- One of his research interests is in the extension of growth curve models, which arise in longitudinal data analysis useful in Biometry and medical science. Another interest is in the SAS-oriented analysis of mixed traditional replication designs useful in industry. Also, He is also interested in mixed linear models, as well as multivariate techniques, especially estimation of chance of miscalculation.

SELECTED PUBLICATIONS

- **Mentz, G. and Kshirsagar, A. M. (2003).** Some Extensions of the Potthoff-Roy Growth Curve Model. Communications in Statistics. .
- **Park, P. and Kshirasagar, A. M. (2002).** A United Method of Analysis of Bio-Assay Designs. Journal of Statistical Studies, Bangladesh 145-157.
- **Shyr, Y. and Kshirsagar, A. M. (1997).** Stepwise Canonical Analysis in Categorical Data. Communications in Statistics 1575-1585.
- **Wang, W. and Kshirsagar, A. M. (1996).** Use of Lattice Square designs in Bioassays. Journal of Biopharmaceutical Statistics 185-99.
- **Park, P. and Kshirsagar, A. M. (1996).** Chance of Misclassification. Australia's

Journal of Statistics 75-83.

- **Weissfeld and Kshirsagar, A. M. (1992).** A Modified Growth Curve Model and its Application in Clinical Studies. Australia's Journal of Statistics 161-168.

PROFESSIONAL AFFILIATIONS

- Fellow, American Statistical Association
- Fellow, Institute of Mathematical Statistics
- Member, International Statistical Institute
- Editorial Board Member, Communications in Statistics

ACKNOWLEDGEMENTS

This brief biographical sketch is prepared ont he basis of an interview taken for Prof. Kshirsagar by prof. M. V. Deshpande, (MS), K.S. Bhanu and (Mrs.) V. Joglekar.

This interview was published as An Evening with **Prof. A. M. Kshirsagar in the journal Gujarat Statistical Review** Vol. 36, Vo. 37, 2009-2010, Nov. 1 & 2. (Pages 59-65) published by GSA.

I thank the authors and the editorial board of GSR for this assistance as well as prof. Kshirsagar by himself.

BOOK REVIEW

ECONOMETRICS BY EXAMPLE

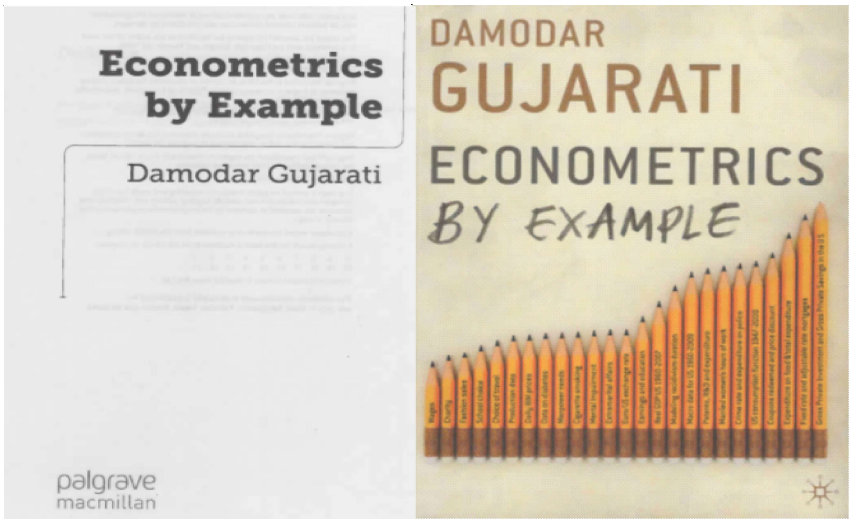
Author DAMODAR GUJARATI
Publisher PALGRAVE, MACMILLA
Edition First 2011
Pages 371

BOOK REVIEW

This is another published book by the author in the lines of econometrics theories and applications. It is presented in 5 parts with appendices and overall 19 chapters.

The book describes basic econometrics models upto the latest approaches with emphasis and discussion in research areas. Here examples means real applications, projects, analysis and interpretations. Extensive use of softwares are shown for different applications with huge and fruitful interpretations.

This is a textbook on research in econometrics fields. The most valuable material presented by the author has some distinct specialities as under.



* Research Mentor, PDPU, Gandhinagar, Email : acbpramukh@hotmail.com
Author is thankful to the referee for reviewing.
(rcd. June '22 / rvd. June '22)

- A wide ranging collection of examples with data on mortgages, exchange rates, charitable giving, fashion sales and more.
- A clear, step-by-step writing style that guides the readers from model formulation to estimation and hypothesis testing, though to post estimation diagnostics.
- Coverage of modern topics such as instrumental variables and panel data.
- Extensive use of Strata and E-views statistical packages with reproductions of output from these packages.
- An appendices discussing the basic concepts of statistics.
- End of chapters summaries, conclusions and excercises to reinforce your learning.

Author has his extensive experience of teaching and research in universities abroad and India. Such a textbook can be a self guide for teachers and research workers in the subject.

Moreover everything is expressed in a very lucid language. Certainly this new book by the learned author is like a treasure presented to the researchers. Each and every library for higher studies must keep this publication invariably in their valuable closets.

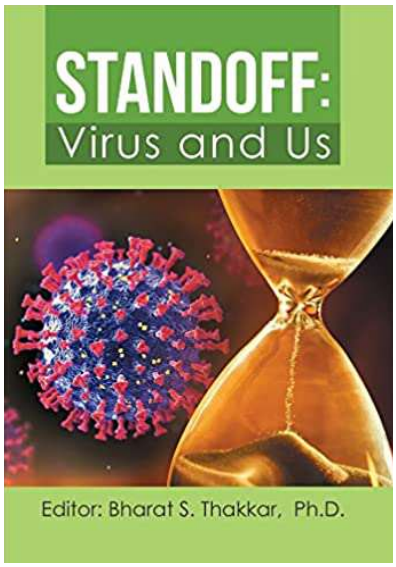
PLEASE REMEMBER THAT THIS IS A BOOK FOR RESEARCH

Ahmedabad
Date : 10 June, 2022

Dr. H.M DIXIT
Statistics Dept.
Arts, Commerce and Science College
PILVAI (N.G.)

STANDOFF: Virus and US

(https://www.xlibris.com/en/bookstore/bookdetails/838230-standoff)



The book published by Xlibris- Chicago based publisher and edited by Prof. Bharat Thakkar in the month of May, 2022 in US is the collective work of 10 authors from different countries. It offers research geared toward understanding the COVID-19 outbreak highlighting the necessity of change management in the development of a comprehensive social media communication strategy in the time of crisis. A pathogen has disrupted the entire planet and its nearly eight billion inhabitants. The book attempts to look at the phenomenon from many different perspectives including how businesses around the world will have to manage this unprecedented, long lasting change in every sphere of society; how businesses and administrations might come forward to minimize the impact of changes resulting from reactions to pandemic.

Various authors of the book try to offer solutions that mitigate or leverage the impact created by the pandemic by employing different change management strategies.

In Chapter 1, the author explains the complexity of the COVID-19 pandemic , the link between pandemic and catastrophic collapse of global systems and change management. Chapter 2 explores what changes and challenges lie ahead for change managers in the current and post-pandemic world. Chapter 3 is an attempt made by the author to study current scenario in the gig economy of the countries like US, UK, Canada, Brazil, Japan and India. It provides a critical evaluation of the efficiency of the change management strategies adopted by these countries like cash payments or stimulus check to gig workers. Chapter 4 describes the economic impact of the

pandemic and provides suggestions for survival and recovery. In Chapter 5, the author presents a collection of global research on how technology start ups and technology entrepreneurship are likely to be impacted by COVID-19. Chapter 6 analyses the application of strategic collaboration and cooperation by the leaders and the scientific community during the COVID -19 pandemic and examines how priorities within the scientific community were established. Chapter 7 reviews the healthcare workers' experiences during the COVID-19 pandemic through case studies and the lessons learned in the preparation for future pandemic of such magnitude. Chapter 8 explains how the virus works and how this knowledge can inform people and allow them to analyze which measures would be appropriate to combat the pandemic. In Chapter 9, the author provides a new approach to change management through quantum computing technology. The author of the Chapter 10 is a renowned journalist and shares journalist's perspective on the interrelatedness of world poverty, global health systems, supply chains and societal balance in a pandemic world.

DR. Bharat Thakkar , the editor of the book and a renowned professor who worked for 5 decades in Chicago area universities , has an uncanny talent for gathering and leveraging the knowledge , experience and insights in leadership and management. He has successfully brought together a team of authors providing their perspectives, predictions and recommendations.

The Foreword is written by Kevin Sorbello , a professor at Capella University, US and full time Manager of Training for a US government contractor.

Book review by **Dr. A. C. Brahmbhatt, Research Mentor, P.D.P.U., Gandhinagar.**

(Special Remarks : In this book, Chapter 3 is wirtten by Dr. A. C. Brahmbhatt. -B B Jani, C.E.)

SV NEWS LETTER

Sanjay G. Raval*

- With this issue of SVJ, June 2022, we complete six months of this year with our Volume No. NSV 18, No.2
- As we all very well know, 29th June of every year is celebrated as statistics day in honour of Prof. P. C. Mahalanobis.

  This year also there is a programme in Gujarat under the auspices of Dept. of Agricultural Statistics, N M C A and Centre for Advanced Agricultural Science and Technology, NAHEP, NAU at Navsari during 29th June 2022 to 30th June 2022 (Two days). The focal theme for the programme is - **Food for Thorught : Applied statistics and its implications.** It is a two days National Seminar and workshop.
- A special programme to celebrate **Statistics Day** is arranged by DES at Gandhinagar on 29th June 2022.
- **Project Announcement**

  Since a long time, we are feeling the lack of awareness of Statistics among graduate and post graduate students specially in Gujarat state. So, we felt to begin some academic awareness program for post graduate students in the Universities of Gujarat state. With the consultation from Professor M. N. Patel and Professor K. Muralidharan, we made a plan to start a project competition in the university. We will request to Professor and Head; and faculty members of the department to make a group of two or four students as one group. If the strength of the students in the corresponding post graduate department is large, then they may form two groups. Professor K. Muralidharan and Professor M. N. Patel jointly will select a title of the project and then we will circulate

* Head, Statistics Dept., Somlalit Commerce College, Ahmedabad - 9
email : drsgraval@gmail.com

the same to each university department. Students will complete the projects, type the project and send it to Professor D. K. Ghosh at his email: ghosh_dkg@rediffmail.com. The group will prepare and finalize the project under any faculty member suggested by Head of the department. These projects will be sent to Professor K. Muralidharan, who consultation with Professor M. N. Patel will arrange for evaluation process through some experts. Final rank will be prepared by Professor K. Muralidharan and will be send to Professor D. K. Ghosh who will declare and announce the result. The awardee students will have to present their project off line/ on line depending upon situations. First prize will be five thousand, and second prize will be two thousand. These prizes will be borne by Professor D. K. Ghosh. Dr. Parag Shah will take care of arranging the on-line project presentation consulting with Professor D. K. Ghosh, Professor M. N. Patel, Professor K. Muralidharan and Professor N. D. Shah.

Project title for the academic year 2022-23 is **"STATISTICAL PREDICTION OF LIFE EXPECTANCY OF INDIAN CITIZEN"**

The last date of submission of complete project is on or before 31 October, 2022.

(Report : Dr. D. K. Ghosh)

OBITUARY

Dr. I. D. Patel

Ishwarbhai Dungardas Patel was born at village Khajdalpur of Visnagar Taluka in Mehsana District of Gujarat. He belonged to a middle class ogricultural family. His mother Kankuben was a loving, simple religious woman. He had two brothers named Lt. Kantilal D. Patel (M.A., B.Ed.) and Prin. Dr. Babubhai D. Patel. (M.Com.) and two sisters named Kantaben and Sitaben. He got education in the local village Khadalpur and then highschool at Uman, which was very far

away from his residence. He got his graduation (B.Sc.) from M. N. Science College, Visnagar with Maths as special subject.

He joined P. G. Dept. of Statistics at Gujarat University and got M.Sc. degree with very good performance. In 1963, he joined Stat. Dept. as demonstrator and gradually he became Lecturer, Reader and Professor in the same department before he retired in 2001.

He had married to a very loving and religious wife Ambaben and he has two sons - (1) Dr. Nilesh I. patel (Eye surgon) and (2) Amish I. Patel (Settled in USA).

I. D. Patel (as we call him I.D.) was a simple, loving, very sincere and hardowrking person. He was very loyal to this subject statistics and teaching work. He did his Ph.D. under world renown Dr. C. G. Khatri in the subject related to Statistical distribution theory and published several research papers. His teaching and research areas were economic statistics, time sereis analysis, sampling techniques, O.R., computer programming, Reliability and life testing etc. He became secretory of GSA established by Dr. Khatri.

His loyalty, simplicity and attachment with Stat. Dept. was unique and remarkable. He had worked in the dept. very sincerely for all the orientation and refresher courses that were given by UGC to this national level department.

He had very serious health problems after retirement in June 2001. Unfortunately I. D. Patel expired suddenly on 1st March 2022.

We all miss him a lot as our loving, sincere and esteemed colleague. May almighty God rest him in peace.

(Report : Prin. M. C. Patel)

Dr. D. N. Shah

Dhirajlal nanchand Shah (commonly named as D. N. Shah) was born in the year 1933 in a small village Mandal, (Nr. Viramgam) in Ahmedabad District of Gujarat State. His father owned a small Tea Hotel in the village,

thus being a middle class person. In childhood Dhiraj lost his father, that was a blow for the family. He received his primary education, higher secondary education in Ahmedabad. He stayed at L. R. Jain boarding in Ahmedabad. He was graduated from Gujarat University and also received his post graduate degree in Statistics subject from Statistics dept., Gujarat University, He remained as a sincere and brilliant student.

He started his acadmeic carrer as a lecturer at Lalan College at Bhuj. Thereafter he also worked at Surat and Billimore in Statistics Department. He joined dept. of Statistics of Sardar Patel University of Vallabh Vidhyanagar, first as a lecturer, then as reader and finally as Professor. He did his Ph.D. under the Veteran Prof. S. M. Shah from S. P. University in 1978. He was retired as professor and head of statistics dept. at S. P. Uni in 1995. His research areas are sampling methods, design of experiments, probability distributions, econometrics etc. He had very good number of research publications in reputed journals. He also guided 4 M.Phil and 4 Ph.D. students.

His wife Vasuben has been very simple, loving and religous woman. She worked for K. G. and Nursery in Vallabh Vidyanagar. Today also the K. G. Institute is very well known as the 'Vasuben Balmandir'. D. N. Shah spent hsi whole life for the developments of his nephew and niece. He was a very religious person. He had done domestic and religious works for Jain religion. He did biggest tapasya known as **Varasi Tap** for a long period. D. N. Shah had very simple life, with no proudy nature. He was very co-ordial and helping to all. He became president of GSA and under his leadership there was a conference of GSA in Bhavnagar. He had ulzimer in 2010 and parkinson in old age Afdtrer life he spent at Mandal and he expiered on 14th May 2022. He had three missions in life teaching, research and religion.

Prof. D. N. Shah will be ever remembered for his simplicity, loving and helping nature. We all miss him a lot. My almighty God rest this soul in peace.

(Report : P. A. Patel)

READERS FORUM

Ashvin J. Patel*

* **Pratheepkumar (Kerala)**
NSV-15, March 2022 issue was remarkable. One very good paper in Applied Maths, also production function estimation was helpful.

* **Bhavin Shah (Indore)**
It is remarkable that anlytical study series is started in SVJ. Articles by Dr. A. C. Brahmbhatt give very useful information. My best wishes for further progress.

* **S. G. Bhimani (Rajkot)**
GSA is performing exemplary works by way of publishing SVJ since last 18 years.

* **H. M. Dixit (Ahmedabad)**
SV News corner, readers forum, biography etc are very interesting features of SVJ along with the routine papers. Please keep it up.

* **Ushakar Gothi (Ahmedabad)**
Data Science is the need of the day. Routine statistical techniques are to be featured with new prospectives. SV team is doing exemplary work. Please keep it up.

* **Rakesh Pandya (Gandhinagar)**
I shall feel happy and obliged if some corner like DATA QUEST can be taken in each issue. Data base research papers may be more fruitful rather than pure theoretical articles. My best wishes.

* **Bhushan J. Bhatt (Bhavnagar)**
Some papers related to computer programming can be inclued in SVJ to make it much more decorative and useful. Nice and exemplary teamwork.

* **B. H. Prajapati (Agri. Uni.)**
I fell that agricultural sector papers are not found. In the whole perpod of 18 years, I think hardly 7-8 papers have come. Please give a booste for this. I appreciate nice team work.

* **P. Mariappan (Trichi)**
Efforts made by SV Team are remarkable. Pl. keept it up.

* **Durai Rajan (Chennai)**
SVJ establishes new trends and traditions. My best wishes. Pl. keep it up.

* Head, Statistics Dept., R. H. Patel Arts & Commerce College, Vadaj, Ahmedabad.
EMAIL : ashvinjpatel@gmail.com

Gujarat Statistical Association

Established : 1969

[Registered under Public Trust Act of 1950 (Bombay)]

R. No. E2502 A'bad - 1974

The objective of the association is primarily to promote statistical ideas in pure and applied fields in the form of study, teaching and research in statistics. The membership of GSA consists of Life / institutional / ordinary members.

Membership	**Fees**	
	In India	Abroad
Institutional Member (for 3 years)	Rs. 2,000/-	US $ 500
Life Member	Rs. 1,500/-	US $ 300
Ordinary Member (p.a.)	Rs. 500/-	US $ 50

Annual Subscription for Gujarat Statistical Review / Sankhya Vignan for individuals, Institutions and library is Rs. 500 (in India), US $ 50 (Abroad)

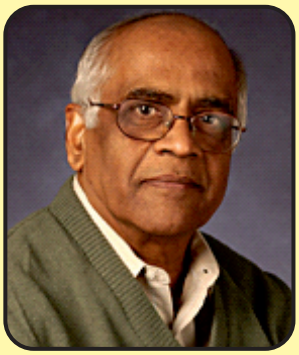
SANKHYA VIGNAN
સંખ્યા વિજ્ઞાન

E-mail : svgsa2015@gmail.com
http://www.sankhyavignan.org

EDITORIAL ADVISARY BOARD

RESEARCH TEAM

PROFESSOR A. M. KSHIRSAGAR*



Professor Anant Madhav Kshirsagar was bon on August 19, 1931 in Sangali (Maharashtra), He received B.Sc. in Mathematics subject in 1948 and M.Sc. in Statistics Subject in 1951 from Bombay University. He received his Ph.D. degree in 1961 under the guidance of **Prof. M. S. Bartlett, FRS in the area of Multivariate Analysis**. During his long academic career standing over more than 5 decads, he taught at **Bombay University** (India), **Manchester** (U.K.), **South Methodist, Texas A and M** (U.S.A) universities and finally **University of Michigan** (U.S.A.) from where he was retired in 2005. He has been an elected fellow of **American Statistical Association** and the **Institute of Mathematical Statistics**. He is also an elected member of the **International Statistical Association.**

He is a recipient of **Don Own Award for excellence in teaching and research**. He is also an author of classical textbook **Multivariate Analysis** and other two books on **Liner Models and Growth Models**.

He has been editor / associate editor of numerous journals including the very famous journal - **Communications in Statistics**.

He has guided more than 35 students for Ph.D. programme from many countries abroad. He was awarded **D.Sc. degree from Manchester University** in 1976. He is also an **Emeritus Professor**. His work areas are **Multivariate Analysis, Design of experiments Growth Curves, Markov Renewal Processes, Response Surfaces, Discriminant Analysis, Liner Models, Bio-statistics and Bio assyas.** He has also worked as visiting professor at Texas A and M University and also at University of PUNE. (Courtesy : GSR, Vol. 36-37, No. 1, 2; 2009-10, Page 59-65)

*** Brief Biographical sketch is given inside the journal.**

Printed Matter **BOOK-POST**

(Journal of GSA, Ahmedabad)

To,

From : **Dr. B. B. Jani**
CE, S. V. Journal
B-14, Bansidhar Apartment, Mirambica Road, Naranpura, Ahmedabad-380013, INDIA
M : +91-9824057902