Strengthening a statistical channel/bridge among researchers and policy makers through statistical collaboration

Amin Muhammad*
Nuclear Institute for Food and Agriculture (NIFA), Peshawar, Pakistan – aminkanju@gmail.com

Eric A. Vance
Laboratory for Interdisciplinary Statistical Analysis (LISA), Department of Applied Mathematics, University of Colorado Boulder, Boulder, USA – Eric.Vance@colorado.edu

Abstract

Pakistan is an agrarian country and the agriculture sector is a backbone of its economy. The majority of the country's employment depends on this sector. This contributes about 19.8% to gross domestic product (GDP). Most of the Pakistanis live in rural areas and their income is related with the development of this sector. This development is extremely reliant on natural factors. Consequently, it is subjected to uncertainty and unpredictability. Moreover, the whole process of agriculture development has weak linkages among its different sectors. The development of this sector has strong relationships with different segments of agriculture, agriculture extension and policy making departments. Unluckily, the progress of this sector remained very low in the past as compared to other agricultural based countries. Most of the issues can be resolved through provision of accurate and timely statistics to policy makers, statistical advice and collaboration to agricultural researchers. There is a need for an effective channel to bridge this sector with other components. Therefore, a comprehensive questionnaire has been designed and distributed among agriculture researchers and policy making officials to know their real issues and gaps and reasons of lacking of inter and intra communications and their other needs. The foremost purpose of this study is to uncover the missing holes and to strengthen the weak linkages with the agricultural research and development to play a role in transfer of real information to the actual stakeholders through a statistical channel/bridge with the help of trained statistical collaborators. The trained statisticians will participate in multidisciplinary agricultural research and will play a key role in bridging the gap between policy makers and others by making them aware of the problems confronted on the basis of the real findings and with the application of statistical models according to the situation.

Keywords: agriculture sector; statistics; accurate information; GDP.

1. Introduction

The economy of Pakistan is agrarian in nature. The agriculture sector (this sector includes sub-sectors namely crops, livestock, forestry and fisheries) is an important element of Pakistan's economy. This sector provides the raw material to down line industries and helps in poverty alleviation impact. The 67.5 percent of the total population belongs to rural area and agriculture is their major source of income (Agriculture census 2010). Agriculture sector contributes 19.8 percent to the gross domestic product (GDP) and provides employment to 43.3 percent of the country's total labour force. Hussain & Khan (2011) studied that there is a positive relationship between GDP and agriculture growth rate and 1 percent increase in the agriculture growth rate can contribute 0.34 percent increase in the GDP. The performance of this sector in the financial year 2016 remained miserable because it gained the growth of -0.19 percent against 2.53 percent growth during the same period in last year (Economic, 2016). This decline can be due to the climatic changes in the world and some internal issues including the access of accurate and timely agriculture statistics to agriculture policy makers and miscommunication among agriculture policy makers, researchers and data collection organizations. Moreover, the researchers have problem to generate quality statistical data due to less knowledge of statistics and communication gap between researchers and statisticians. Similarly, Mughal (2016) identified twenty eight problems of this sector under Techno-Economic, Natural, Socio-Economic and Financial problems.

Ahmed (2000) analyzed the agricultural policy in Pakistan and said that there are many problems in formulation of fresh or incorporating amendments in the existing agricultural policies. There is a need to deal with these issues systematically and scientifically. Some unreliable and spotted statements on agricultural policy has been observed if we look into the publish reports in the last few years in Pakistan. During the meeting with agricultural policy makers and researchers regarding the development process of questionnaires for this study, it has been perceived that the officials responsible to formulate or amendment/upgrade the policies related to agriculture have no access to the accurate and timely data. The coordination between the researchers and policy makers is very weak during the formulation of policies.

Amin & Vance (2016) identified the issues related to agricultural researchers while designing their experiments, collecting data, analysing data and communication of results to the researchers (non-statisticians). They studied that there is lacking of statistical facilities in the agricultural research institutes. There is a gap between researchers and statisticians. They proposed to create a statistical collaboration laboratory namely Pakistan Laboratory for Statistical Collaboration (PLASCO) at Nuclear Institute for Food and Agriculture (NIFA), Peshawar, Pakistan a subsidiary institute of Pakistan Atomic Energy Commission to fill this gap. The overall vision of PLASCO was to promote effective collaborations between statisticians and agricultural researchers that lead to improvements in agricultural productivity in Pakistan. This idea of developing a statistical collaboration laboratories in the agricultural research system was proposed by Msemo and Vance (2015). Moreover, Goshu, A. T. (2016) studied that creation of statistical collaboration laboratories can improve the quality of research through provision of proper statistical collaboration to researchers (non-statisticians).

In this paper, we included a very important stakeholder (Agricultural policy makers) and proposed a modified approach that can also play a very crucial role to increase the agricultural productivity that ultimately can improve the incomes of poor farmers. This paper is still in draft form because we are currently collecting the data from agricultural policy makers to help us identify and construct useful theory to help create a statistical channel/bridge among researchers and policy makers through statistical collaboration.

A new theory is proposed to strengthen the statistical channel among researchers, policy makers and farmers in Sections 2. Research methodology and study design is explained in Section 3-4. Section 5 contains the case studies that are conducted to test the proposed theory. The overall conclusion is provided in Section 6.

2. Proposed Theory

Statistics is quite often a stumbling block preventing impactful research from being conducted and disseminated. There are two reasons that lack of statistical knowledge impedes the conduct of research.

2.1. Theoretical Reason # 1

The researchers do not even try to conduct an innovative study because they know they will not be able to design the study or experiment appropriately or that they will not be able to analyze the data correctly once collected. This reason prevents statistically sophisticated research from even being attempted, which is a problem because cutting-edge research nowadays is often statistically sophisticated.

2.2 Theoretical Reason # 2

Several things can go wrong without proper statistics. We have identified three such constructs. Construct 1 is a poorly designed study so that the data are insufficient to support an important result (this is like a Type III error in which the data answer the wrong question or that the right question cannot be answered with the data that were collected). Construct 2 is that the statistical analyses are too basic and more modern or advanced methods will lead to a stronger conclusion (this is like a Type II error, or

a false negative, in which a true result is rejected, often because of lack of power because of a statistical method poorly suited to the data and study design). Construct 3 is that the statistics are done incorrectly and lead to conclusions not actually supported by the evidence (this is like a Type I error, a false positive in which the reported result/conclusion is actually false but reported as true).

Statistics is often a stumbling block preventing impactful research from being disseminated and used. One mechanism explaining this is that to be published in a widely read and cited journal, a paper must employ proper statistical methods. An experiment or study may have been conducted well, but if the results are not communicated effectively, high quality journals will not publish them and they will not be widely read. Another related issue is language. If a researcher's native tongue is not English, he or she is at a disadvantage when submitting to an English-language journal. Also, if the results (especially the statistics) are not communicated in a way for non-experts to understand, then a published paper might not be understood by a wide audience and therefore its potentially important results might be ignored or overlooked.

The typical researcher is an expert in his or her field, but not an expert in statistics. Therefore, many will suffer the problems with statistics and statistical communication described above. So, statistics is very important to conduct and communicate high impact research.

Proposed solution that will not work: train every researcher to become an expert in statistics. This would require an estimated X hours of training from Y statistical training and education experts. Currently in Pakistan there are only an estimated Z statistical training and education experts, and relatively few of these will be effective teaching such a wide variety of researchers (for example, a PhD statistician may have an interest in economics and would be able to teach economists enough statistics to become experts in economic statistics, but this PhD statistician might fail to effectively teach soil scientists to become experts in the statistics related to soil science because of his or her lack of experience with soil science). Our proposed solution is to train statisticians to communicate and collaborate with researchers. An estimated T hours will be required to train U statisticians to communicate and collaborate with non-statisticians. These U statisticians have already spent an estimated V hours becoming experts in statistics. Then the U statisticians can work for W hours on S projects to provide the lacking statistical expertise through statistical collaboration. These hours (T + W) are much less than X and therefore are a much more efficient way to enable and advance research through expertise in statistics.

Similarly, lack of statistics knowledge and skill prevents policy makers from incorporating evidence from well-designed research studies. And similarly, we propose creating statistical collaboration laboratories to bridge the gaps between statistics and policy, between statistics and research, and between research and policy.

3. Research Methodology

The testing of above cited theory is planned in Pakistan with the following procedure.

1. How many researchers (what percent, estimated) do not follow-up on ideas because of Theoretical Reason #1 of lacking sufficient statistical expertise to even get started? Is the primary issue not knowing how to design a study? How to design a proper experiment? Analyzing the data after collection? Collecting the data? Managing the data (maybe it's a Big Data problem and the researcher lacks Big Data skills)?
2. How common is Construct 1? Construct 2? Construct 3?
3. Are there other "constructs" that lead to a lack of high quality research because of lack of statistical skills and knowledge?

4. How common is Theoretical Reason #2 lack of effective communication of statistics? How often (estimated) is high quality research not published in high quality journals because of statistics or language?
5. Estimates for S, T, U, V, W, X, Y, and Z. How much less is (T + W) than X? Does this effect persist over time? What if the U statisticians also taught statistical short courses and workshops based on their knowledge of what exactly is needed in the fields they work in? Over time how would that time investment of teaching specific short courses pay off?
6. What are the stumbling blocks for policy makers to not use statistics?
7. What are the stumbling blocks for policy makers to not incorporate relevant research?
8. Do policy makers believe that collaborating with statisticians will help improve policy?

4. Study design and region

The first author of this paper had several meetings with researchers and policy makers to know their real problems/issues related to statistics and get guidance to design an effective and comprehensive questionnaires for testing the theoretical concepts (1-8) mentioned in Section 3. The policy making officials and the researchers are target populations of this study. The detailed questionnaires are designed and distributed among the respondents through personal visits and emails. The questionnaire included quantitative and qualitative questions. The data collection is ongoing. The database will be established in MS Excel which will include primary data from responses to the questionnaire and calculated weights of each criterion for various groups of experts/researchers. The data will be analyzed and the statistical models will be developed to fill the gap between the researchers and policy makers and to recommend statistical techniques for improving of their statistical knowledge.

5. Modified theory

After completion of the survey study, we will present the results. The modified theory will be developed based on those results. The following three case studies will be conducted to support the proposed modified theory.

Case study #1: How statistical collaboration enabled or accelerated a specific research project in Pakistan.

Case study #2: How statistical collaboration with a policy maker led to a policy decision based on data.

Case study #3: How statistical collaboration bridged the gap between research and policy.

6. Conclusions

We will test our proposed theory with completion of survey study. We are optimistic that the model of PLASCO and LISA 2020 will enable and accelerate research and policy making and bridge the gaps between statistics, research, and policy. PLASCO and LISA 2020 needs to be financially supported and its model scaled up to create more stat labs across Pakistan, South Asia, and the world.

References

Ahmed, S & Roger P. M. (2000). Agricultural Policy Analysis in Pakistan: Illustrations in the use of the Policy Analysis Matrix, Centre for Management and Economic Research, Lahore University of Management Sciences, Pakistan.

Economic, S. (2016). Pakistan Economic Survey 2015-16. Islamabad: Economic Adviser's Wing, Fianance Division, Government of Pakistan.

Goshu, A. T. (2016). Strengthening Statistics Graduate Programs with Statistical Collaboration-The Case of Hawassa University, Ethiopia. International Journal of Higher Education, 5(3), p217.

Hussain, A., & Khan, A. Q. (2011). Relationship between agriculture and GDP growth rates in Pakistan: An Econometric Analysis (1961-2007). Academic Research International, 322-326.

Msemo, E., & Vance, E. A. (2015). LISA 2020: Impacting Agricultural Productivity in Tanzania through the Wheels of Statistics. Proceedings of the International Statistical Institute's 60th World Statistics Congress.

Mughal, N. (2016). Agriculture. Retrieved from CSS Forum: http://www.cssforum.com.pk

PBS. (2010). Agricultural Census 2010 - Pakistan Report. Retrieved from Pakistan Bureau of Statistics: http://www.pbs.gov.pk