Proceedings of the 4th International Conference on Statistics: Theory and Applications (ICSTA'22)
Prague, Czech Republic – July 28- 30, 2022
Paper No. 163
DOI: 10.11159/icst22.163

Leveraged Study Design for Identifying Dominant Causes of Variation

Mahsa Panahi, Stefan H. Steiner, Jeroen de Mast
Department of Statistics and Actuarial Science, University of Waterloo
Waterloo, Canada, N2L 3G1
mpanahi@uwaterloo.ca; shsteiner@uwaterloo.ca; jdemast@uwaterloo.ca

Extended Abstract

Excessive variation in critical to quality process outputs is a common challenge in manufacturing industries. For variation reduction, most process quality improvement (variation reduction) frameworks follow Juran's diagnostic and remedial journeys [1], that is, first using some methods to find the cause(s) of output variation (the diagnosis) and then, seeking a solution for eliminating the effect of the identified cause(s) (the remedy).

Among all causes of variation, usually only a few have a big impact on the overall variability [2]. Shainin refers to them as the dominant cause(s) [3, 4]. Finding the dominant cause(s) requires a systematic strategy. The Shainin System™ [3, 5] is a coherent statistical stepwise variation reduction strategy with several problem-solving techniques. One of the techniques associated with the Shainin System™ that aims to help identifying the suspect dominant causes is *group comparison*, which exploits the concept of leveraging by comparing the extreme parts [5]. To do so, we select two groups of six or more (typically eight) parts, one group consisting of parts with large and the other with low quality characteristic Y values. Then, only for these selected parts, we measure as many suspect dominant cause input characteristic X's as possible. If X is a dominant cause, its values must be substantially different between the two groups. Shainin [3] suggests using the Tukey end-count test [6] to compare the X values in the two groups.

Although the investigation plan based on leveraging is an efficient way of gathering information in searching for a dominant cause using relatively small sample size, the Shainin analysis procedure is less than ideal. The following provides some of the most critical critiques. First, the end-count method does not implement well for discrete inputs (since ranking order for binary inputs has many ties and it is not unclear how to extend it for categorical inputs with three or more levels). A more substantial criticism is that the Shainin group comparison procedure inefficiently uses the data. The reason is that it ignores all the Y values for the non-extreme parts. Moreover, even for the parts in the extreme groups, it discards the observed Y values to group membership information. The most important point of critique, however, concerns reliability. It frames the problem as a hypothesis test (whether a candidate X affects Y) instead of an estimation problem (whether X is a dominant cause of variation in Y). As a result, the end-count-based analysis, which is expected to identify dominant causes of output variation, is more likely to find minor causes. We provide a simulation study that quantifies our claim.

To overcome the above issues, we propose a new analysis method that frames the question as an estimation problem based on maximum likelihood. We critically assess the efficiency of the new procedure by comparing different study designs via simulation. Moreover, we conduct another simulation study demonstrating that our proposed analysis procedure is considerably more reliable than Shainin's procedure in identifying the dominant cause(s) of output variation. Finally, we provide a tangible example of how our superior proposed method works.

References

[1] J. M. Juran and F. M. J. Gryna, *Quality Planning and Analysis* (2nd ed.). New York: McGraw-Hill Book, 1980.
[2] F. M. Gryna and J. M. Juran. *Juran's Quality Control Handbook* (4th ed.). New York: McGraw-Hill, 1988.
[3] R. D. Shainin, "Strategies for technical problem solving," *Quality Engineering*, vol. 5, no. 3, pp. 433-448, 1993.
[4] R. D. Shainin, "Statistical engineering: six decades of improved process and systems performance," *Quality Engineering*, vol. 24, no. 2, pp. 171-183, 2012.
[5] K. R. Bhote and A. K. Bhote, *World Class Quality* (2nd ed.). New York: American Management Association, 2000.
[6] J. W. Tukey, "A quick, compact, two-sample test to duckworth's specifications," *Technometrics*, vol. 1, no. 1, pp. 31–48, 1959.