Holy Hollerith’s holes and another stories in Statistical Computing

Jorge A. Navarro Alberto
Visiting professor (Universidad Autonoma de Yucatan)

Abstract

In an interesting paper published in the ASA’s website, David Alan Grier gives us an account on how statistical computing began during the 1920s and 1930s, when US universities and research labs began to acquire the early IBM mechanical punched card tabulators. It is well known that these punched cards had been invented by Herman Hollerith for the 1890 U.S. Census and that Hollerith formed the Hollerith Tabulating Machine Company to manufacture and market these devices. This company was later merged with two other firms in 1911 to form the Computing Tabulating and Recording Company or C.T.R. In 1924, C.T.R. was renamed International Business Machines (IBM). In the 1920’s, universities used these machines not only for tabulating and computing summary statistics but also for fitting more complicated statistical models such as analyses of variance and linear regressions.

This talk starts considering the “Prehistoric ages” of computers and computing like those illustrated above, and then goes through a discussion on how statistical computing, data analyses, graphical user’s interfaces, and another computing concepts have been changed during the course of our “Computing Era”. Then we raise two questions: 1) Which advantages and disadvantages we have had in this open field that computers and software represent, where entertainment and “real computing purposes” coexist? and 2) How graphical interfaces help / prevent us to make most out of the computing potential found in computers? We will bring up these matters for discussion, and finally I will give my opinion on the way statistical computing has been taught to our 21th century students, especially for those outside Computing Science. A particular example on how to cope with the overwhelming and widespread use of R in Science, will be given.