Gaines, B. R. (1970). Interaction and significance. Behavioural Technology 1, 11-12.

Interaction and Significance

For the experimenter/data-analyst computers have so far been a mixed blessing. The capability of inverting a 50 x 50 matrix not only knocked the bottom out of the market for factor-analysis-based Ph.Ds (the record being a 62 x 62 matrix taking seven years to invert (a later analysis on a computer took three minutes (what is more, the earlier analysis proved correct))), but also meant that there was no excuse for not gathering sufficient data to generate a 50 x 50 matrix. No longer was there sympathy for the psychologist who experimented with four subjects (himself and three rats), and extrapolated his results to the entire domain of animal life – overnight the voracious computer increased the data required for respectability by at least two orders of magnitude.

This much is history and accounts for the increase in population of albino rats, and the swing towards sociology, where the rules allow diversity to be swamped by numbers. However, only those who have cried through the dark nights of desperate data analysis, immersed in hundreds of yards of line-printer paper, can appreciate the truly insidious impact of a computer – when we want one correlation coefficient we get one hundred, when we want ten we get ten thousand. Out they come in neat formation across the enormously wide paper (it is no answer to narrow the paper – it has been tried).

At one time this plethora of numbers was welcome, but then the dark truth dawned – if we have one hundred coefficients any one needs to be one hundred times as significant to be publishable! Whereas, previously we subsisted at the 5% level, now only the .05% level (unattainable dream) is acceptable. To combat this effect the size of experimental populations was increased 10-fold, then a 100-fold. This increase in effort should have swamped the generosity of the computer, but no – two other effects appeared which between them wiped out any advantage of the increased experimental population.

The first is statistical, although undreamed of by the statisticians in their ivory towers surrounded by normal populations of albino rats with Gaussian tails; as the experimental population increases so does the scarcity of additional members until finally in desperation one is adding individuals whose normalcy is virtually zero – this has been formalised by statisticians as the 'Law of Large Numbers'. The second effect is political, a viper in the bosom of any ivory tower; experiments with larger populations take a longer time but fill only the same number of papers. This may be overcome by measuring a proportionately larger set of variables in the same population. However, the number of cross correlations between the variables goes up as a square of the number measured and, returning to our previous discussion, so does the significance level required of each individual correlation. It has been shown that the overall effect is of experimental populations asymptotic to infinity, which will, for practical purposes, reach there by the year 2,060 (by this time it is also reliably reported that the world population will have reached a point where each individual has .1 square metre of space available to him). This is known by statisticians as the 'Law of Very Large Numbers'.

Although the interactive computer is not yet a means to populations control in general (this will form the topic of a further article) it has come in time to have a massive impact on data analysis, amounting virtually to the repealing of the Laws of Large Numbers. One example will suffice to demonstrate the enormous savings possible through the use of a computer terminal. Consider that you have measured some ten variables within a population and are considering the cross-correlations between them. With a conventional computer system you would receive a sheet of paper with a hundred correlation coefficients on it and would examine them for significance, requiring 40 times the overall significance for each individual result (anybody who can see why 40 should not be reading this article (anybody who can see why not 40 will not be reading this article)).

Suppose now you have an interactive terminal and analyse the same data; the results now remain within the computer and, since it is not examining them, their significance does not deteriorate with their number – it only decreases as we ask for each result. Thus, if we request only one correlation coefficient to be printed the significance level required will be, let us say, 5%. If we ask for five results to be printed, the significance of each one will be required to be 1%, and so it goes on – it is clear that the more we know, the less we have to ask, and the less work we have to do to obtain a given level of significance. This is eminently commonsensical, but statisticians have given it no name – I suggest that it be called 'The Law of Diminishing Satisfaction', since it is the essence of voyeurism that increased peeping has decreasing effects.

This is but one example – many new and exciting statistical techniques based on human interaction with the computer are being discovered every day. A logical extension of the technique described above is to have the computer print out only those correlation coefficients which exceed a certain value. Since the computer itself is an entirely electro-mechanical device, its selection cannot be said to be subjective and the increased significance obtained in this way is certainly worth the extra programming effort required. For those of us whose correlation coefficients never exceed .2 anyway, even these techniques may not be powerful enough and there are others which go far beyond them: Tukey's vacuum cleaner techniques for cleaning up data have their logical extensions and enable virtually any experiment to lead to significant results.

The statistical referee of this paper noted, 'that the logical conclusion is not to analyse the data'. This seemed so obvious that I failed to mention it – the concept antedates Aristotle – *O si sic omnia!*

#