CONTINGENCY TABLES

Prof. E. B. Wilson
Harvard University

In biological and medical experiments many contingency tables arise that cannot be analyzed by the chi-square test because of low cell frequencies. A method for treating such cases is presented. Although illustrated with one four-fold universe with two marginal totals fixed, besides the total number in the sample, the general principle can be stated for v cellular universes with not necessarily equal numbers of cells, and with L totals remaining fixed, including the v totals of the size of the subsample from each universe.

Consider now the sample of N from a four-fold universe having two characters A and B, with probabilities $p_1 = p_{AB}$, $p_2 = p_{\alpha B}$, $p_3 = p_{A\beta}$ and $p_4 = p_{\alpha\beta}$. If n_i is the number in a sample of N having the attribute associated with p_i, then the probability of observing n_1, n_2, n_3, n_4 is given by

$$P = \frac{N!}{n_1!n_2!n_3!n_4!} p_1^{n_1}p_2^{n_2}p_3^{n_3}p_4^{n_4}.$$

We restrict our further attention to those tables that have the following totals fixed: $n_1 + n_3 = (A)$, $n_1 + n_2 = (B)$ and $N = n_1 + n_2 + n_3 + n_4$, and also satisfy the condition that the probability of their occurrence shall not vary from table to table by virtue of the values of p_1, p_2, p_3 and p_4.

These conditions are sufficient to determine an associated universe that represents the appropriate null hypothesis. It can also be shown that the probability of an observed table arising from a universe satisfying the null hypothesis is

$$\left[\sum \frac{1}{n_1!n_2!n_3!n_4!}\right]^{-1} \frac{1}{n_1!n_2!n_3!n_4!},$$

where the summation is over all samples which could arise, satisfying the fixed totals.

The rule can now be made that the significance of a table is to be determined by the sum of the probabilities of the table and of all other tables no more probable.

The condition that the probabilities shall not vary from table to table, and the rule just stated, will give a test of significance.

Statistical Flowers Caught in Amber

Paul A. Samuelson

Since I remember well the war-time MIT seminars in statistics now being reproduced in abstract form, I am happy to accept the editors' invitation to reminisce about those times.

Chance alone turned up these Abstracts in the University of Chicago libraries. Although it was my secretary (and Harold Freeman's), Eleanor Prescott Clemence, who typed up these mathematical abstracts, all of us had forgotten they were ever compiled. With probability not minute, Harold Freeman would have sent a copy of them to our friend W. Allen Wallis, who with certainty approaching unity throws away nothing. (The initials W. A. W. on the manuscript Stephen Stigler stumbled upon in the Chicago archives are in the unmistakable schoolboy hand of the Honorable W. Allen Wallis.)

Paul A. Samuelson is Institute Professor Emeritus, Massachusetts Institute of Technology, E52-383C, Cambridge, Massachusetts 02139.

Actually, with faculty blessings, this seminar series was conceived and executed by two graduate students: Lawrence Klein, who was to become MIT's first Ph.D. in Economics and our first home-grown Nobel Laureate; and Joseph Ullman, then studying economics but in the course of the war's windup in Europe later to be enticed into a career in mathematics by Gabor Szegö. Laurie and Joe both as introducers of the speakers; Harold Freeman and I would both cringe and delight in the unpredictable algebraic felicities of their unrehearsed introductions. (Sample: when the illustrious Richard von Mises was to be presented, his many fames as a pioneer had not run ahead of him; so our student impresario left it at, "Although I don't know why, our speaker is supposed to be a very famous scholar.")

It is amazing that, in this epoch after Pearl Harbor, when faculty was dispersing to various war-time labs and graduate student bodies were shrinking to a small core of transients and women, two active students could still attract without stipends so brilliant a group of speakers. Most were