One percent software error rate too optimistic?

Editor:
Since reading an article in last year's issue on program testing (''Hints on Test Data Selection . . .'' DeMillo, et al., Computer, April 1978), I have been interested in program error statistics as they may apply to software testing. The article quoted a frequency distribution of errors by type of HOL statement involved, taken from a PhD thesis by E. A. Youngs.1

Dr. Youngs, who is now with Bell Telephone Laboratories, has made available his original data and procedures, from which it appears that he may have done the most extensive study to date on software errors and their correlates. Among others, the following consequences of Youngs' data appear to be of interest:

1. His discrete distribution of errors by statement type is a conditional distribution, i.e., it provides values of

\[ P(j|x), \quad j=1, 2, \ldots 14, \]

where \( x \) represents the set of "erroneous" instructions, and \( j \) is the class of statement with which the error is associated in Youngs' scheme of 14 classes.

2. The 1258 "errors" on which the distribution is based are not original indigenious errors but contain repetitions due to multiple "runs" and errors introduced by "fixes."

3. If 60 programs in the study are examined for mean percent error \( E \), the percent of total source statements represented by the "first-run" errors (these should be close to the original errors) is 10.5 percent, with a sample standard deviation of about 8.5 percent.

4. A linear regression analysis of percent of source statements in error \( E \) on the number \( N \) of programs writ-

\[ E = -0.076N + 12.6 \]

and a similar regression of \( E \) on programmers' prior years experience \( Y \) yields

\[ E = -0.053Y + 11.9 \]

(5) A histogram of the 60 sample points from item 3 above reveals a unimodal sample distribution reminiscent of a chi-square or exponential distribution.

These results seem to indicate that the "magic number" of one percent frequently mentioned or inferred for error rate2,3,4 may be too optimistic, although error rate obviously depends on programmer capability.

Two measures related to this are number of years programming experience and number of programs written. For Youngs' study, programmers had written an average of 28.6 programs and had an average of 2.7 years experience, though both of these had large standard deviations (21.6 for number of programs and 2.9 for years experience).

G. B. Hawthorne, Jr.
Warner Robins, GA