DOCUMENT RESUME

ED 106 335 TM 004 450

AUTHOR Reilly, Richard R.
TITLE Estimating Total Score and Item Completion Statistics When All Examin es Do Not Finish.
PUB DATE [Apr 75]
NOTE 12p.; Paper presented at the Annual Meetings of the American Educational Research Association and the National Council on Measurement in Education (Washington, D.C., March 30-April 3, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS *Correlation; *Prediction; *Scores; Statistical Analysis; Testing; *Timed Tests

ABSTRACT

For most tests administered with time limits some examinees complete all items while others do not. It is often useful to know what the distribution of items completed would be if the number of items on the test were much larger. It may also be of interest to estimate the correlation between working speed, as measured by the number of items completed, and the total score. The method described assumes that working speed and total score are bivariate normally distributed in the population of interest and allows estimation of the means and standard deviations for items completed and total score, and the correlation between items completed and total score given a theoretically infinite number of items in the same time limit. (Author)

Estimating Total Score and Item Completion Statistics When All Examinees Do Not Finish

Richard R. Reilly

Educational Testing Service

Princeton, N. J.

NCME

Washington, D. C.

AERA #14.16
NCME #15

Estimating Total Score and Item Completion Statistics

When All Examinees Do Not Finish

Richard R. Reilly
Educational Testing Service

There are situations in which it is useful to know the extent to which examinees' working speed is related to their total score. If no examinees finish the test than the correlation between the number of items completed (c) and total score (x) should provide a good estimate of this relationship. In many cases, however, a significant proportion of examinees finish the test and all receive the same score on c , thereby making the usual correlation coefficient inappropriate. Ideally, a much longer test (long enough so that no examinees finished) composed of additional parallel items could be readministered with the same time limits to a new sample and the desired estimate obtained. Because this approach is rarely feasible an alternative method is proposed which can be used to estimate r_{xc} , and at the same time provide the following additional parameter estimates for a theoretical test consisting of a large number of parallel items:

$\bar{c}$ = the mean number of items completed.

s_c = the standard deviation of items completed.

s_x = the total score standard deviation.

$\bar{x}$ = the total score mean.

Estimating the Mean and Standard Deviation of Items Completed

The method involves the assumptions that items completed and total score have a bivariate normal distribution and that examinee working speed would not be affected by an additional set of parallel items. Figure 1 illustrates the observed and theoretical distribution of items completed for a test which a large proportion of examinees

Insert Figure 1 About Here

did not finish. The distribution has been divided into two segments: A, the actual distribution of items completed; and B, the theoretical distribution for examinees who finished the test. The two segments are divided by a line at c_o which represents the actual number of items in the test.

To obtain the estimate s_x the following relationships are employed:

$$p_n = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_c} e^{-z^2/2} \, d_z \qquad (1)$$

$$\mu_1' = \frac{1}{p_n\sqrt{2\pi}} \int_{-\infty}^{z_c} z \, e^{-z^2/2} \, d_z \qquad (2)$$

$$\mu_2' = \frac{1}{p_n\sqrt{2\pi}} \int_{-\infty}^{z_c} z^2 \, e^{-z^2/2} \, d_z \qquad (3)$$

$$\sigma' = (\mu_2' - (\mu_1')^2)^{1/2} \qquad (4)$$

$$s_c = s_c'/\sigma' \qquad (5)$$

where

p_n is the proportion of examinees not finishing the test.

z_c is the unit normal deviated corresponding to p_n .

μ_1' is the first moment of the truncated unit normal distribution with terminus z_c .

μ_2' is the second moment of the truncated unit normal distribution.

σ' is the standard deviation of the truncated normal distribution.

s_c' is the estimated standard deviation of items completed for those not finishing the test.

s_c is the estimate of the standard deviation of items completed for the complete theoretical distribution (i.e., finishers and non-finishers with an infinite number of additional parallel items).

Once s_c has been obtained it becomes possible to obtain $\bar{c}$, the estimated mean number of items completed. Actually it is possible to obtain $\bar{c}$ in two different ways as follows:

$$\bar{c}_1 = c_o - z_c s_c \quad ; \text{ or} \tag{6}$$

$$\bar{c}_2 = \bar{c}' - s_c \mu_1' \tag{7}$$

where c_o is k-1 for a k-item test, and $\bar{c}'$ is the mean number of items completed for examinees not finishing the test. If the distribution is normal these two estimates should agree fairly closely, if the two estimates differ widely the distribution of non-finishers should be examined using a method similar to that proposed by Donlon (1975). It may be that the normality assumption does not hold.

Table 1 shows the mean and standard deviations for truncated normal distributions with truncation points corresponding to different proportions of non-finishers, p_n. Table 1 is presented for illustrative purposes rather than as a complete table. Values of μ and σ can be computed for exact values of p_n with the help of a complete table of the normal distribution.

Insert Table 1 About Here

Estimating the Correlation Between Items Completed and Total Score

Given s_c, the following formula can be used to obtain r_{xc}.

$$r_{xc} = \frac{s_c\, r'_{xc}}{\sqrt{s_c^2\,(r'_{xc})^2 + (s'_c)^2\,(1 - (r'_{xc})^2)}} \tag{8}$$

where

r_{xc} = the correlation between items completed and total score for non-finishers.

Estimating the Mean and Standard Deviation of Total Score

The mean of the total score distribution can be estimated as follows:

$$\bar{x} = \bar{x}' - b(\bar{c}' - \bar{c}) \tag{9}$$

where

$$b = r'_{xc} \frac{s'_x}{s'_c}$$

s'_x = total score standard deviation for non-finishers, and

$\bar{x}'$ is the mean total score for non-finishers.

The standard deviation of the total score can be estimated as

$$s_x = s'_x \sqrt{1 - (r'_{cx})^2 + (r'_{cx})^2/(\sigma')^2} \quad . \tag{10}$$

The reader will recognize formulae (8) and (10) as variants of those presented by Gulliksen (1950, pp. 137-138) in his discussion of explicit selection. Variants of (8), (9), and (10) were derived by Cohen (1955) as maximum likelihood estimates of parameters of the bivariate normal distribution when one variable is restricted. Cohen (1950) also presented an alternative method of estimating the mean and standard deviation of a sample drawn from a truncated normal distribution. The method requires the computation of

$$v_1 = \sum_{i=1}^{N_n} (c_i - c_o)/N_n \quad ; \text{ and} \tag{11}$$

$$v_2 = \sum_{i=1}^{N_n} (c_i - c_o)^2/N_n \quad . \tag{12}$$

These values are then substituted in formulas provided by Cohen (1950) and an iterative procedure must then be used to arrive at estimates of the mean and standard deviation of c . Cohen's procedures are computationally more laborious than the previous method but have the advantage of providing maximum likelihood estimates of the mean and standard deviation.

Example: A new 30-item test of arithmetic is administered to 1,000 students. Total time allowed for the test is 20 minutes. The test is completed by 45% of the students. The test developers are interested in adding more items to the test but would like to know what the relationship between working speed and total score is. They are also interested in knowing what the mean and standard deviation of items completed would be for a much longer test. Finally, they want some idea of the total score mean and standard deviation for a much longer test.

In this particular case 550 students did not complete the test. The first row of Table 2 shows statistics based on the 550 non-finishers only. The second row of Table 2 shows estimates for the total sample which were obtained using the relationship given above.

Insert Table 2 About Here

Comment

The methods presented here are based on assumptions which should be more widely investigated with empirical data. Good empirical validation of the utility of these assumptions and the methods presented here could be done with large samples of examinees and tests which no (or an insignificant percentage of) examinees finish.

References

Cohen, A. C. Restriction and selection in samples from bivariate normal distributions. Annals of Mathematical Statistics, 1955, 26, 884-893.

Cohen, A. C. Estimating the mean and variance of normal populations from singly truncated and doubly truncated samples. Annals of Mathematical Statistics, 1950, 21, 557-569.

Donlon, T. Estimating rate-of-work param ers in the assessment of test speededness. Paper presented at NCME, Washington, D. C., 1975.

Gulliksen, H. Theory of mental tests. New York: John Wiley & Sons, 1950, 137-138.

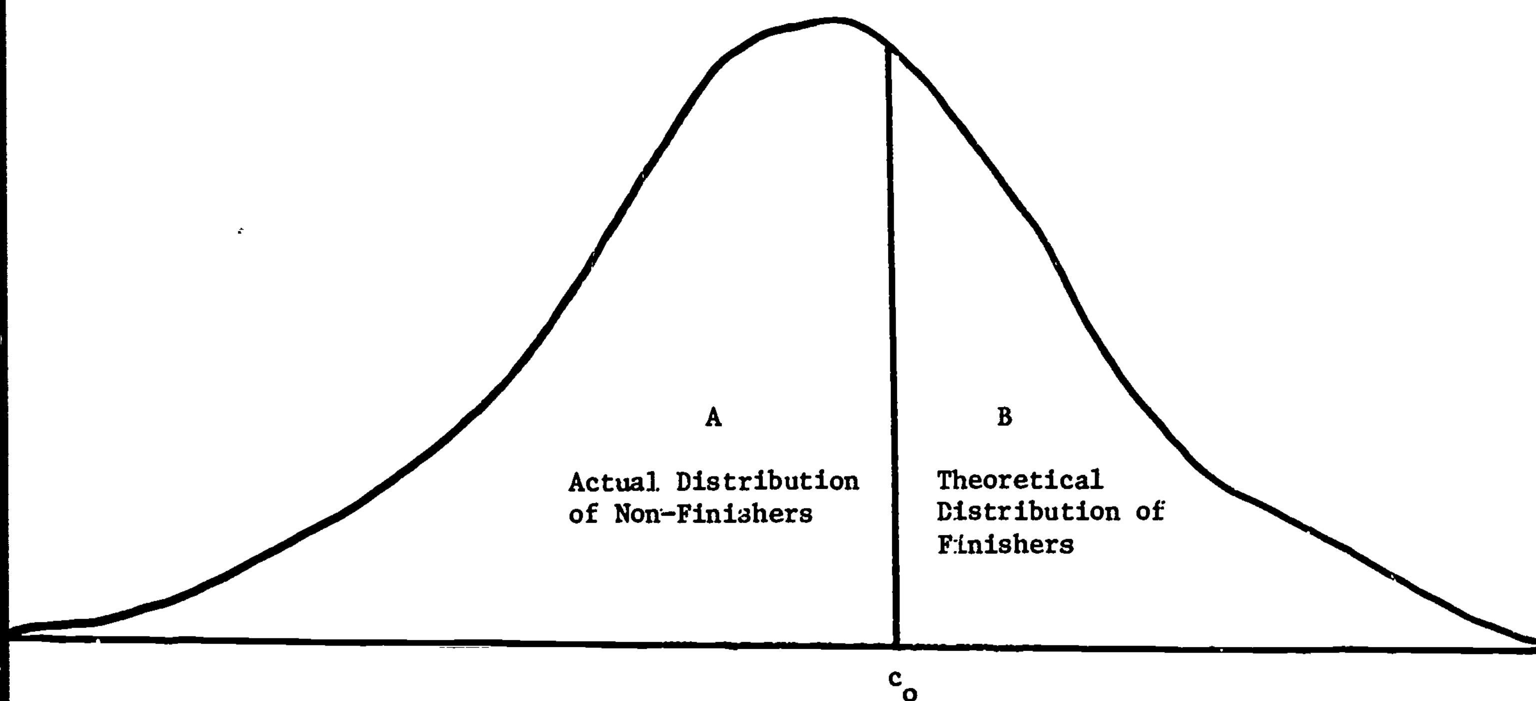

Figure 1

Illustration of Actual and Theoretical

Distribution of Items Completed

Table 1

Means and Standard Deviations for the Unit Normal Distribution

Truncated at Points Corresponding to p_n

p_n	μ'	σ'
.05	-2.06271	.34400
.10	-1.75498	.40431
.15	-1.55439	.43936
.20	-1.39981	.47774
.25	-1.27111	.50021
.30	-1.15898	.51878
.35	-1.05828	.54111
.40	-0.96586	.55471
.45	-0.87957	.58367
.50	-0.79788	.60281
.55	-0.71965	.62584
.60	-0.64390	.64981
.65	-0.56984	.67506
.70	-0.49670	.70201
.75	-0.42370	.73123
.80	-0.34995	.76355
.85	-0.27430	.80029
.90	-0.19500	.84385
.95	-0.10856	.89980

Table 2

Estimates of Parameters for Non-Finishers and

Theoretical Estimates for All Examinces

	Mean # Items Completed	S.D. Items Completed	Mean Total Score	S.D. Total Score	Correlation Between Items Completed and Total Score
Computed Estimates for Non-Finishers	23.60	4	16	3.00	.30
Theoretical Estimates for All Examinees	28.20	6.39	17.04	3.20	.45