DETERMINATION OF LIKELIHOOD OF BELONGINGNESS OF AN INDIVIDUAL TO A NORM GROUP BY MEANS OF STANDARD SCORES ON INDEPENDENT PSYCHOLOGICAL MEASURES

Ajit Kumar Mukherjee, Corpus Christi State School
Texas Department of Mental Health and Mental Retardation

In psychology it is always a problem to determine with some certainty whether an individual belongs to a group or not. An example may make the issue clear. Suppose that Mr. A takes a test composed of various subtests to determine A's suitability for a training program. The test is designed to measure various independent abilities needed to complete the training program successfully. The test A has taken has been standardized using a sample of people who completed the training program successfully (norm group). There are three possibilities:

1. A's abilities may be above the norm group.
2. A's abilities may be like the other members of the norm group.
3. A's abilities may be below the norm group.

The question is often asked --- Does A belong to the norm group? Cronbach and Glesser (1953) recommended a model for assessing similarity between profiles designed to handle the question "How similar is Person 1 to Group Y?" Cronbach and Glesser - D^2 and Mahalanobis - D^2 are identical to each other when variates are standardized and uncorrelated (Cronbach and Glesser, 1953). However, for studying groups of persons, Cronbach and Glesser - D^2 did not prove to be very effective. In 1928, Pearson recommended "coefficient of racial likeness" which was designed to measure the similarity between two groups or the similarity of an individual to a group. Unfortunately, Pearson's index proved unsatisfactory. Since Pearson, several techniques have been recommended by several individuals. Cattell (1949) introduced the concept of r_p as a coefficient of pattern similarity. One of the assumptions of Cattell was that for computing r_p, variates were needed to be uncorrelated. Williams (1969) in his study found that moderately correlated variates could be used effectively for computing r_p.

When statistics only deal with probability, all statistical models deal with 'chance'. And, no statistical model is designed to predict certainty.

In this article, an attempt has been made to deal with the issue of chance of an individual to belong to a group.

PROCEDURE

The following theorems have been used to develop the technique:

1. If the random variable X is $N(\mu,\sigma^2)$, $\sigma^2>0$ then the random variable $V = (X-\mu)^2/\sigma^2$ is χ^2 with df = 1.

2. Let $X_1, X_2 \ldots . X_n$ denote a random sample of size n from a distribution which is $N(\mu,\sigma^2)$. The random variable $Y = \sum_{i=1}^{n} \frac{(X_i-\mu)^2}{\sigma^2}$ has a chi-square distribution with n degree of freedom.

3. $F = \frac{U/r_1}{V/r_2}$ has F distribution when U and V are independent chi-square variables with r_1, and r_2 degrees of freedom respectively.

MODEL

We have a test with K subtest such that each subtest measures independent psychological trait of an individual. A sample of n individuals are randomly selected to standardize the test. On each independent subtest standard score (Z-Score) of each n individual is calculated. Then we have a matrix,

	1	2	3	K
1	Z_{11}	Z_{12}	Z_{13}	Z_{1k}
2	Z_{21}	Z_{22}	Z_{23}	Z_{2k}
3	"	"	"	"
"	"	"	"	"
"	"	"	"	"
"	"	"	"	"
"	"	"	"	"
n	Z_n	Z_n	Z_n	Z_{nk}

squaring each Z - Score we have,

Z^2_{11}	Z^2_{12}	Z^2_{13}	Z^2_{1k}
Z^2_{21}	Z^2_{22}	Z^2_{23}	Z^2_{2k}
"	"	"	"
Z^2_{n1}	Z^2_{n2}	Z^2_{n3}	Z^2_{nk}

Now, Z^2_{ij} is chi-square with one degree of freedom. Since, each subtest is independent, $\sum_{i=1}^{k} Z^2_{ij}$ is chi-square with k degrees of freedom. Let that be written as χ^2_{ik}.

Since, performance of each individual is independent of each other, $\sum_{i=1}^{n} i\chi^2_k$ is chi-square with n·k degrees of freedom.

Now, let us consider a case who is not a member of the sample used for standardization and whose Z-scores are $Z_{m_1}, Z_{m_2} \cdot \cdot \cdot \cdot Z_{m_k}$

$\sum_{i=1}^{n} {}_i\chi^2_k$ and ${}_m\chi^2_k$ are independent chi-squares.

A F-ratio can be obtained and F ratio can be defined as:

$$F = \frac{\sum_{i=1}^{n} {}_i\chi^2_k / n \cdot k}{{}_m\chi^2_k / k}$$

A F-table can be utilized to determine the significance of F with degrees of freedom as nk and k at .05 or .01 level. The obtain F ratio indicates to which extent between the subjects variability is larger than that of within subject variability. When F is significant, one may state that a significant positive correlation exists. Therefore, it may be concluded that there is a likelihood that the subject outside the "norm group" belongs to the norm group.

ILLUSTRATION

Let us consider a sample where n=4 and a test with three subtests (k=3). The following matrix represents standard score of each subject in each subtest.

S_1	-1.1	1.0	0.8
S_2	1.2	-1.3	0.9
S_3	1.4	2.0	1.3
S_4	2.0	1.7	-0.5

Square of Standard Score Matrix:

1.21	1	0.64	${}_1\chi^2_3 = 2.85$
1.44	1.69	0.81	${}_2\chi^2_3 = 3.94$
1.96	4.0	1.69	${}_3\chi^2_3 = 7.65$
4.00	2.89	0.25	${}_4\chi^2_3 = 7.14$

$\chi^2_{12} = 2.85 + 3.94 + 7.65 + 7.14 = 21.58$

$\chi^2_{12}/_{12} = \frac{21.58}{12} = 1.79$

Let us consider two individuals with standard scores in the same test but not members of the "norm group".

S_5	1	0.5	2
	1	0.25	4

$\therefore \chi^2_3 = 5.25$

$\chi^2_3/_3 = \frac{5.25}{3} = 1.75$

What is the likelihood that S_5 belongs to the norm group?

$F = \frac{1.79}{1.75} = 1.02$ F with degrees of freedom 12 and 3, to be significant at .05 level needs to be 8.74 F is not significant. Likelihood of S_5's belongingness is not significant.

Let us consider:

S_6	0.3	0.5	0.2
	0.09	0.25	0.04

$\chi^2_3 = 0.38 \qquad \chi^2_3/_3 = 0.13$

$F = \frac{1.79}{0.13} = 13.84$

Here, F is significant at $P < .05$ level.

Likelihood of S_6's belongingness to the norm group is significant.

DISCUSSION

Data from a research project (Title IVC, funded by Texas Education Agency) was used to determine the limitation of the model. A sample of N=600 with measures on seven independent variates was used. In some cases, variates were moderately correlated (Table 1).

	1	2	3	4	5	6	7
1	1	.33	.32	0.06	.36	.28	.13
2		1	.35	0.02	.37	.34	.14
3			1	0.07	.33	.28	.13
4				1	.10	.11	.07
5					1	.32	.16
6						1	0.10
7							1

Various random samples of different sizes and different number of variates were drawn from the original sample (N=600) to examine the limitation of the model. The criterion for belongingness to the norm group was the agreement between mental age of an individual and the mean mental age of the norm group. That means:

(a) mean mental age for each random sample was computed from the raw scores (refer to the report of the research project), and

(b) an individual was selected at random and the model discussed in this article was used to compute the F-ratio. When F-ratio was significant, mental age of the individual in question was computed. If the mental age of the individual and the mean mental age of the group were in agreement(range ± 0.5 years), prediction of the likelihood of belongingness to the norm group was considered to be accurate. The following are the inferences:

(1) At least measures on five independent variates and a sample size of 30 were needed for predicting the belongingness of an individual to the "norm group".

(2) N = 50 and above tend to predict the belongingness of individuals with very flat profiles (${}_{m}\chi^2_k \leq 3.52$, # variates = 7, N = 50)

(3) The model was most suitable for samples when N = 40 and # of variates = 7.

REFERENCES

Cattell, R.B., r_p and Other Coefficients of Pattern Similarity, Psychometrika, 1949, 14, 279-298.

Cronbach, L.J. and Glesser, G.C., Assessing Similarity Between Profiles, Psychological Bulletin, 1953, 50, 456-472.

Hogg, R.V. and Craig, A.T., Introduction to Mathematical Statistics (3rd). New York, Macmillan Company, 1971.

McNemar, Q., Psychological Statistics (4th). New York, John Wiley and Sons, Inc., 1969.

Pearson, K., On the Coefficient of Racial Likeness, Biometrika, 1928, 18, 105-117.

Williams, J.M., An Empirical Study of Cattell's Coefficient of Profile Similarity, Unpublished doctoral dissertation, University of Alabama, 1969.

SPECIAL REPORT

Mukherjee, A.K., Ander, S., Sluyter, G.V., and Leal, A., Measurement of Intellectual Potential in Mexican-American School Age Children. Corpus Christi State School, Texas Department of Mental Health and Mental Retardation, June, 1976.