DOCUMENT RESUME

ED 077 964 TM 002 773

AUTHOR	Reilly, Richard R.
TITLE	A Note on Minority Group Test Bias Studies.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RB-73-15
PUB DATE	Feb 73
NOTE	11p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Comparative Analysis; *Minority Groups; Prediction; Statistical Analysis; Technical Reports; *Test Bias; Test Interpretation

ABSTRACT

Comparisons of majority and minority group regression lines for purposes of assessing test bias may, under certain conditions, be viewed as comparisons of conditional bivariate distributions. Where these conditions hold, findings should reveal parallel regression lines except for a special case. Given the conditions described, one implication is that even when the test is a parallel form of the criterion, lines with equal slopes but unequal intercepts should be found. (Author)

RB-73-15

RESEARCH BULLETIN

A NOTE ON MINORITY GROUP TEST BIAS STUDIES

Richard R. Reilly

This Bulletin is a draft for interoffice circulation. Corrections and suggestions for revision are solicited. The Bulletin should not be cited as a reference without the specific permission of the author. It is automatically superseded upon formal publication of the material.

Educational Testing Service
Princeton, New Jersey
February 1973

A NOTE ON MINORITY GROUP TEST BIAS STUDIES

Richard R. Reilly

Educational Testing Service

Abstract

Comparisons of majority and minority group regression lines for purposes of assessing test bias may, under certain conditions, be viewed as comparisons of conditional bivariate distributions. Where these conditions hold, findings should reveal parallel regression lines except for a special case. Given the conditions described, one implication is that even when the test is a parallel form of the criterion, lines with equal slopes but unequal intercepts should be found.

A NOTE ON MINORITY GROUP TEST BIAS STUDIES

Richard R. Reilly

Educational Testing Service

In recent years a fair number of studies have attempted to assess the cultural bias in tests by conducting validity studies for minority and majority groups. In the academic context the criterion is usually grades, while in the industrial situation supervisor's ratings are often utilized. Typically, such studies compare a minority group regression line with a majority group regression line. The procedure described by Gulliksen and Wilks (1950) may then be used to test sequentially differences between (a) the standard errors of estimates, (b) the slopes, and (c) the intercepts. Another method might be to use the F distribution to test equality of slopes and intercepts (e.g., Cleary, 1968). Failure to reject all null hypotheses would suggest that the two groups were drawn from the same bivariate normal population.

It is the purpose of this note to suggest that if certain conditions hold, identifying groups for the purposes of predictive bias studies may result in a crude form of statistical conditioning, which makes it likely that regression lines with equal slopes but unequal intercepts will be found.

Though most cultural bias studies have dealt with categorically defined groups, if we assume a continuous normally distributed variable underlying the categorizations, then it is possible to consider a trivariate normal distribution of the following variables all expressed in standard score terms ($\mu = 0$; $\sigma = 1$):

y_1 , a test variable

y_2 , a criterion variable

y_3 , a variable underlying, or highly correlated with, the group categorizations, which will be referred to as a sociocultural variable in this paper.

The density function can be written as:

$$f(y_1 , y_2 , y_3) = \frac{|R|^{\frac{1}{2}}}{(2\pi)^{\frac{3}{2}}} e^{-\frac{1}{2} Y'RY} ,$$

where R is the inverse of the variance-covariance (or, in this case, correlation) matrix,

$$V = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix} ,$$

and Y is the random vector with elements y_1 , y_2 , y_3 .

If we now take the conditional bivariate density function of y_1 and y_2 for some given value of $y_3 = Y_3$, we find

$$f(y_1 , y_2 | Y_3) = \frac{|R^*|^{\frac{1}{2}}}{(2\pi)} e^{-\frac{1}{2} Y^{*\prime}R^*Y^*}$$

where R* is now the inverse of the variance-covariance matrix V* . V* may be written in terms of the elements of the original V matrix as

$$V^* = \begin{bmatrix} (1 - \rho_{13}^2) & (\rho_{12} - \rho_{13}\rho_{23}) \\ (\rho_{12} - \rho_{13}\rho_{23}) & (1 - \rho_{23}^2) \end{bmatrix}$$

and the elements of Y* may be written as

$$Y^* = \begin{bmatrix} y_1 - \rho_{13}Y_3 \\ y_2 - \rho_{23}Y_3 \end{bmatrix} ;$$

in other words, we now have a bivariate normal distribution with means $\rho_{13}Y_3$ and $\rho_{23}Y_3$.

We can now find the slope of y_2 on y_1 as follows:

$$\beta_{y_2/y_1} = \frac{(\rho_{12} - \rho_{13}\rho_{23})}{(1 - \rho_{13}^2)} .$$

Note that this slope, which is a conditional slope, will be the same no matter what value of y_3 is chosen.

Next we find the intercept,

$$\kappa = \rho_{23}Y_3 - \beta_{y_2/y_1}\,\rho_{13}Y_3$$

$$= \frac{Y_3(\rho_{23} - \rho_{12}\rho_{13})}{(1 - \rho_{13}^2)} .$$

Thus, although the slope will remain invariant for different values of y_3 it can be seen that the intercept will vary depending on the value of y_3 chosen, except for the special case of $\rho_{23} = \rho_{12}\rho_{13}$.

Under certain conditions, then, predictive bias studies might be viewed as comparisons of conditional normal bivariate distributions. This conditioning may occur as a result of the identification of group categories which represent crude points on a third normally distributed variable, y_3.

If one accepts the trivariate normal assumptions, then it is already clear that the slopes will be the same for all groups and that intercepts will differ (except for the special case of $\rho_{23} = \rho_{12}\rho_{13}$). Correlations may differ within groups because of selection on the predictor, which may be more extreme for some groups than others. This does not affect the above argument, however, since the slopes, intercepts and standard errors should not be affected by selection.

The special case where $\rho_{23} = \rho_{12}\rho_{13}$ will result in equal intercepts no matter what values of y_3 are chosen. Three specific instances of this set of conditions are worth noting:

(a) The case where y_3 is uncorrelated with both test and criterion (i.e., $\rho_{13} = \rho_{23} = 0$).

(b) The case where the test and criterion are perfectly correlated (i.e., $\rho_{12} = 1.0$).

(c) The case where y_3 is more highly correlated with the test than with the criterion to the extent that the condition $\rho_{23} = \rho_{12}\rho_{13}$ holds.

Although all three examples would result in equal regression lines only, cases (a) and (b) can be considered completely unbiased. Case (c), as Thorndike (1971) has pointed out, is "unfair" in terms of the group proportions which would be selected given a fixed cutting score on the predictor.

Whether the lines differ and how they differ will depend on the elements of R^* and the values of y_3 used to form the conditional bivariate distributions. In general, however, it is safe to say the following about most recent studies in this area:

(a) A positive correlation exists between test and criterion.

(b) If the sociocultural variable is defined such that individuals generally recognized as more advantaged have higher scores, the sociocultural variable is positively correlated with both test and criterion.

Consider now the situation which is most desirable (in terms of the bias problem) short of having a perfect test-criterion correlation. This would be a situation where the test is actually a parallel form of the criterion. That

is, except for uncorrelated errors of measurement the test is measuring the same thing as the criterion. We have, therefore, a bivariate normal population in which ρ_{12}, the test-criterion correlation, is actually a reliability coefficient. Here, clearly, $\rho_{13} = \rho_{23}$, since y_1 and y_2 are parallel and $\rho_{13} = \rho_{23} \leq \rho_{12}$. The situation where all three correlations are equal is highly unlikely since this would mean the sociocultural variable, y_3, was parallel with y_1 and y_2. Hence, we can say with some generality that $\rho_{13} = \rho_{23} < \rho_{12}$.[1]

If we now find two conditional normal bivariate density functions for y_1 and y_2 by taking two points on y_3, say, $Y_{3(A)}$ and $Y_{3(B)}$, where $Y_{3(A)} > Y_{3(B)}$, we can deduce the following:

(1) The slopes of the regression of y_2 on y_1 will be the same for both conditional distributions since the covariance matrix V^* remains invariant for all values of y_3.

(2) The intercepts for the conditional distributions will be

$$\kappa_A = \frac{Y_{3(A)}(\rho_{23} - \rho_{12}\rho_{13})}{(1 - \rho_{13}^2)} \text{, for group A}$$

and

$$\kappa_B = \frac{Y_{3(B)}(\rho_{23} - \rho_{12}\rho_{13})}{(1 - \rho_{13}^2)} \text{, for group B} \quad .$$

Since all correlations are positive and $\rho_{23} = \rho_{13}$, we find that $\kappa_A > \kappa_B$.

That is, the intercept for the "high" sociocultural group is greater than the intercept for the "low" sociocultural group, or to put it another way, group B is consistently overpredicted if predictions are based on the group A regression equation. It is also clear that as $\rho_{12} \rightarrow 1.0$ the difference in intercepts

becomes smaller. Though the above argument is given for the case where y_1 and y_2 are parallel test forms the same results will be obtained in the case where y_3 is correlated with factors common to y_1 and y_2 and uncorrelated with all other factors. It is worth noting that in this situation if the socio-cultural variable is ignored and the group A regression line is used, any cutting score on the predictor will yield the "correct" proportions of each group in terms of the percentages expected to perform at a corresponding level on the criterion.

Although a number of studies have reported parallel regression lines (e.g., Cleary, 1968; Temp, 1971) a recent report by Campbell, Pike, and Flaugher (1969) is particularly worth noting because the criterion employed was a carefully constructed, objective, job-knowledge test. For seven of eight predictors studied the regression lines for minority and majority groups were parallel with lower intercepts for the culturally disadvantaged group.

Further empirical evidence for this phenomenon is presented in Table 1.

Insert Table 1 about here

The "criterion" in this instance is a roughly parallel form of the "predictor." Both variables are separately timed reading comprehension sections of the Law School Admission Test (data were collected during a regular national administration). Here, it can be seen that although the slopes are very nearly the same, the difference in intercepts is marked, with the lower intercept belonging to the regression line for Blacks.

The specific conditions described in this paper are not the only way in which overprediction can occur, of course. Linn and Werts (1971) describe some more general possibilities in a recent paper. The argument presented here assumes that groups are defined by conditioning on the same bivariate

distribution. Although it may well be that the groups are from quite different bivariate populations, it seems reasonable that when objective measures such as tests serve as both predictors and criteria, the argument will hold.

Conclusions

When conducting minority group bias studies investigators should recognize that in some cases they may be comparing two conditional bivariate distributions from the same general bivariate population. Under the conditions described when the sociocultural variable is correlated with factors common to the predictor and the criterion but uncorrelated with all other factors the finding of "over prediction" for the lower sociocultural group should be expected.

References

Campbell, J. T., Pike, L. W., & Flaugher, R. F. Prediction of job performance for Negro and White medical technicians. Project Report 69-6. Princeton, N. J.: Educational Testing Service, 1969.

Cleary, T. A. Test bias: Prediction of grades of Negro and White students in integrated colleges. Journal of Educational Measurement, 1968, 5, 115-124.

Gulliksen, H., & Wilks, S. S. Regression tests for several samples. Psychometrika, 1950, 15, 91-114.

Linn, R. L., & Werts, C. E. Considerations for studies of test bias. Journal of Educational Measurement, 1971, 8, 1-4.

Temp, G. Test bias: Validity of the SAT for blacks and whites in thirteen integrated institutions. Journal of Educational Measurement, 1971, 8, 245-251.

Thorndike, R. L. Concepts of culture fairness. Journal of Educational Measurement, 1971, 8, 63-70.

Footnote

[1]This statement is not completely general since it does not take into account cases where y_1 is the y_2 true score (or vice versa) or where y_3 is the underlying true score. I am indebted to Dr. Robert L. Linn for pointing this out.

Table 1

Selected Statistics Computed Separately for Blacks and Whites[a]

	N	$\bar{X}p$	$\bar{X}c$	Sp	Sc	r_{pc}	Slope	Intercept
Blacks	500	12.3	11.5	4.2	3.9	.61	[illegible]7	4.5
Whites	500	15.4	14.8	4.5	4.2	.62	[illegible]	5.9

[a]The criterion, $\bar{X}c$, is a separately timed reading comprehension section of the LSAT, as was the predictor $\bar{X}p$.