

CHAPTER VIII

ACADEMIC PREDICTION IN GENERAL CONCLUDING REMARKS, SUMMARY OF RESULTS AND SUGGESTIONS

So far, the results of the discriminant analyses have been derived on the usual assumption of the normality of the distribution of scores. Welch has given the solution for the discriminant problem in a more general way. The only requirement (according as Welch has shown) is the probability ratio $\frac{p_1}{p_2} > K$, where K is suitably chosen. We therefore examine such a solution empirically for the data of SSCE marks in question.

The following birariate table (Figure 5) shows the 'totals' and 'failures' in each cell. The lines show total frequencies, while crosses show fail-frequencies in each cell. From these, the empirical probabilities of passing p_1 in each cell, are calculated and shown for extreme positions at the corners (p_2 being $= 1 - p_1$).

By inspection, we find that barring a few cases (corresponding to a small chance error, (say 5%), the region of acceptance and rejection and the doubtful region can be given as shown by dotted lines, in regard to the two predictors SSCE English and Mathematics.

SSCE - ENGLISH MARKS

85-89														
80-84									I 1.0					
75-79						I +		I 1.0	I 1.0	II 1.0	II 1.0		II 1.0	
70-74				III I	II			I 1.0	II 1.0	III 1.0	II 1.0	III 1.0	II 1.0	
65-69		I + 0.0	I +	III +	IIII +	I +	IIII ++	IIII ++	II 1.0	IIII + 0.88	I 1.0		I 1.0	
60-64		I +	II +	IIII ++	IIII +	IIII ++	IIII ++	IIII ++	IIII + 0.83	IIII 1.0	II 1.0	III 1.0	II 1.0	
55-59	IIII ++ 0.0		IIII ++ 0.0	IIII ++	IIII ++	IIII ++	IIII ++	IIII +	IIII + 0.60	IIII 1.0	IIII 1.0	IIII 1.0	II 1.0	
50-54	II ++ 0.0	I +	IIII ++ 0.0	IIII ++	IIII ++	IIII ++	IIII ++	IIII ++	IIII 1.0	I + 0.0	II + 0.5			
45-49			II +	I + 0.0	I + 0.0		IIII ++	IIII ++	II 1.0		II 1.0			
40-44	I + 0.0		I + 0.0	IIII ++	I +	IIII ++	I + 0.0		I 1.0	I 1.0				
35-39	I + 0.0			II ++ 0.0	II ++ 0.0	I +	IIII ++	IIII ++	IIII 0.0		I + 0.0	I + 0.0		
	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95-99	

SSCE - MATHEMATICS MARKS

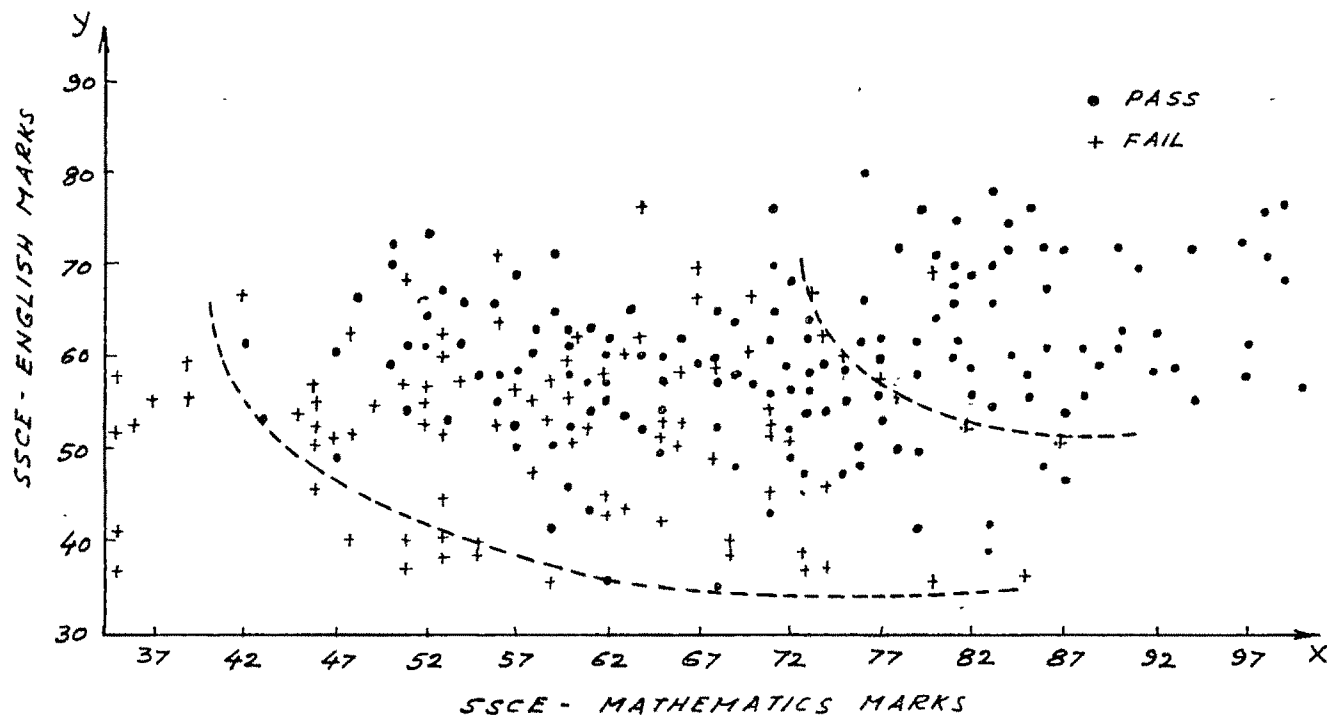
BIVARITE TABLE SHOWING EMPIRICAL PROBABILITIESAND REGIONS OF ACCEPTANCE & REJECTIONFIG. 5

More exactly, the empirical solution can be obtained from scatter-diagram of passes and failures, which specify the regions of acceptance and rejection. We thus have, as an illustration, the following plan (Fig. 6) empirically derived correspondingly to a small chance error of zero and .01 ($\alpha_2=0$; $\alpha_1=.01$) for the two predictors.

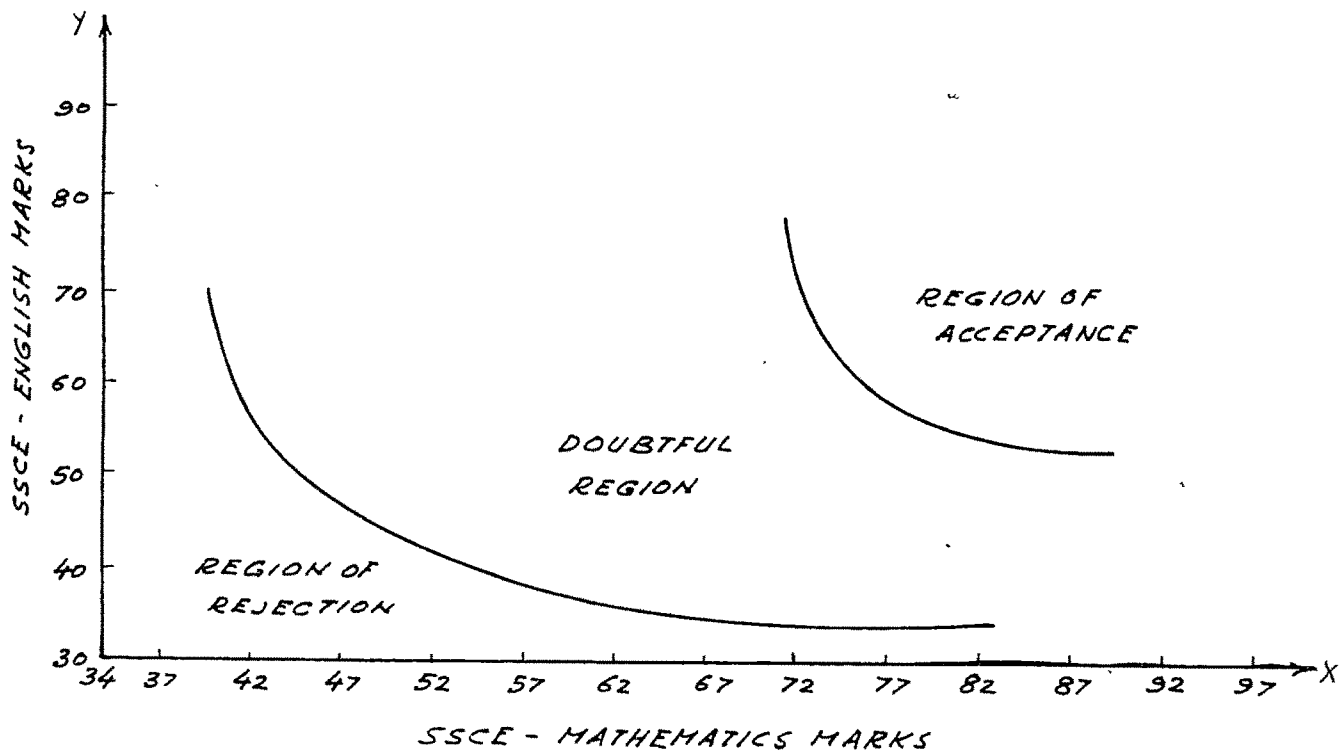
The region of acceptance covers all such cases as should be accepted straightway for admission. It should be the policy of everyone concerned to see that all such deserving students of great college promises are admitted to the course, if they so desire. In the institutions of higher learning which maintain high standards of education, the cases in this region should be admitted.

The doubtful region represents such cases for whom decision cannot be taken at the desired level (here .01) of chance error on the basis of the two given predictors and hence should be pursued for further scrutiny with additional test/tests. If we can afford to have one or more aptitude tests these should be availed of and suitable method be gainfully used.

The region of rejection covers all such cases which hardly has any chance of success and hence should be rejected, if the two predictors in question serve the only



SCATTER DIAGRAM OF PASSES AND FAILS



AN EMPIRICAL PLAN CORRESPONDING TO A SMALL CHANCE
 ERROR OF ZERO OR .01 (i.e. ONE PERCENT)

FIG. 6

criteria for admission. In many situations we find the selections are based on these two important predictors of college success, namely English and Mathematics. If , however, we have some additional aptitude test score, the cases in the rejection region be re-examined alongwith the cases in the doubtful region, eventhough they stand little or no chance of being selected on further scrutiny.

We find that when the admission procedures are liberal and no hard and fast rule is prevalent for admissions, as in the present data, the failures occurred as 103 out of 278 total cases admitted, that is nearly 37 percent. This is type I error while type II error is zero. But the above analysis clearly shows that 17 undeserving cases could have been rightly rejected, and this would have reduced type I error to 32.9 percent, type II error still being zero. As bloom asserts, reduction of even 1 percent means saving of thousands of rupees. Thus we find that the results are better than if not used at all.

Further as we leave the rejection boundary, the error of second kind which is zero goes on increasing and error of first kind goes on decreasing. This happens upto certain point, until error of second kind becomes equal to the error of first kind and then afterwards the error of se-

cond kind becomes greater than the error of first kind and finally the error of second kind reaches its maximum value while the error of first kind tends to zero.

What does the error of I kind imply?

The error of I kind implies wastage or college risk. Due to this error, cases predicted pass actually fails. The error of II kind implies individual or personal risk. Due to this error cases predicted 'fail' actually pass. The optimum solution offered by the data, for which type I error equals type II error strikes balance between these two types of risks and gives a certain percentage of correct classification. This is the maximum level of accuracy that can be obtained with the given predictors. We observed in chapter V, that the maximum level of accuracy in terms of efficiency of correct classification is approximately 75 percent with regard to the two predictors-SSCE English and Mathematics. When maximum attainable level of accuracy is reached, the prediction can be said to have been saturated with regard to these predictors. We found that further improvement could be made only by addition of one or more reliable and valid tests and that too, by resorting to the method underlined in chapter VI. Thus we had the maximum level of accuracy, with regard to SSCE English, Mathematics and

General Science in terms of efficiency of correct classification as 77 percent approximately, which we could not obtain by usual procedure.

This study was taken up with the conviction that the academic success of students largely depends on their scholastic abilities. To start with, the investigation was confined to the two important factors of academic success, namely English and Mathematics. This is equivalent to starting with a null hypothesis for the effects of other factors to be null or zeroes. When the maximum level of accuracy with regard to the two predictors of English and Mathematics is reached, the influence of other factors, if any, could then be studied, one after another by suitable methods and if found significant, further improvement in prediction could be effected by the method given in chapter VI, and thus further extraction and utilization of factors could be achieved.

LIMITATIONS

Limitations of Educational Measurements

The main purpose of this thesis was to examine into the statistical aspect of the problem, that is, to seek answer to the pertinent question-whether academic prediction could be improved by some statistical method, if we have the data which has reached the upper level of accuracy

(correlation of .70 or more). Consistent with the purpose, suitable data was selected and taken up for study. Naturally, this empirical investigation has its limitations of data and though the methods would be applicable in general, similar results could be expected only under stable system of examination and identical situations of test data and instruments. For any instrument, it cannot be said that it is the last word as far as prediction of the given criterion is concerned. Better results could be obtained if we have better tools or instruments. As for example, in the present cases, we find that the points corresponding to passes and fails are intricately mixed up along the discriminant line. It is here that the need for items properly discriminating between passes and fails is felt. If we have more and more number of items which can give maximum discrimination along this line, we can have better results in terms of efficiency of correct classification.

Limitations of Statistical Methods

We must recognize, the limitations of statistical methods at this stage of discussion also. No statistical method can increase the accuracy of the given data. In other words, it is not possible, with any statistical method, to improve the data which is initially inaccurate

due to poor test or examination. Secondly, though the statistical methods are very important and without them sound generalizations cannot be made, they are applicable to the group as a whole, under study.

The prediction problem can undergo resinement in three stages:

- 1) Instruments of intellectual abilities
- 2) Statistical methods
- 3) Personal factors or non-intellectual factors

The results achieved through stages one and two, that is, with a given instrument and with any statistical analysis will no doubt be applicable to the group as a whole, but in some individual cases which can be considered as chance errors, other non-intellectual factors such as, health, personality and socio-economic condition etc. etc. might influence the college performance, and hence cent percent prediction is seldom possible to acquire.

Efforts to account for non-intellectual factors systematically have not been encouraging. Bhatnagar, as also Bloom, reports very low correlations of these factors with the academic success. If, however, we could have powerful inventory or personality scales, their in-

fluence can be tested, and if found significant, due allowance for them could be made. When maximum attainable accuracy by all means for the first two stages is reached the concept of overachieving and under-achieving bears meaning. Only then the studies can be taken up for extracting non-intellectual background factors.

SUMMARY OF RESULTS AND SUGGESTIONS

The major objective of this thesis was to see if academic prediction could be improved by some method over the usual linear regression approach. The findings of this investigation can be conclusively put as follows :

- 1) The prediction of college mathematics is significantly improved by nonlinear quadratic regression in case of SSCE achievement test in mathematics as well as numerical aptitude test. In view of this, item validity indices for mathematics may be examined in modified form which can take into account the quadratic component.
- 2) In case of the prediction of college mathematics from the combination of English and numerical tests, the curvilinear relationship given by the general equation of the second degree raises the level of accuracy of the prediction significantly well above the linear combination.

Performance in college mathematics is very important for college success in a science course. By and large, failures at college examination in Science occur due to failure in college mathematics. As such, the above results are significant in view of their implications in validating test results against the criterion of college mathematics.

3) The method of selection given in chapter VI for making use of additional test/tests gives improved results, and hence be used to achieve further gains in prediction.

4) The clear empirical evidence as to upper and lower curvilinear solutions can be appropriately utilized for selection of top cases and rejection of bottom cases in suitable situations. They provide better results than if not used at all.

We know that 100 percent accurate prediction is not possible with any instrument or with any method; inherent complex nature of higher mental processes and the personal chance factors do not permit such an accurate measurement of psychological and educational measurements. However, it is realized that with a given instrument, there is a maximum attainable accuracy that can be reached by a suitable method, and our main objective was to explore this method and the maximum level of accuracy. It is hoped that, over

and above what has been achieved through this work, it will open new vistas for further work in the realm of academic prediction.

The Report of Education Commission (1964-66)
remarks-

'The search for good and reliable methods of selection is one of the important problems in higher education and vigorous research is needed to evolve them. Even in advanced countries, satisfactory techniques of selection have not been developed as yet. While search for good methods goes on, we have to begin the programme with such ad hoc methods as are available.'