Liu, C.-C. et al. (Eds.) (2014). Proceedings of the 22nd International Conference on Computers in Education. Japan: Asia-Pacific Society for Computers in Education

Study on Pass-Fail Prediction for the National Radiological Technologist Examination by Discriminant Analysis Using Mock Examinations

Hiroyoshi WATANABE[a*], Shigeru SASAKI [b], Shinsuke AIKO[b] & Shigemi SAKAMOTO[c]
[a] *Learning Technology Laboratory, Teikyo University, Japan*
[b] *Graduate School, Teikyo University, Japan*
[c] *Nihon Institute of Medical Science, Japan*
*hiro@ics.teikyo-u.ac.jp

Abstract: The objective of this study was to establish a technique for pass-fail prediction for a national examination from the results of intramural mock examinations, for practical use within a school. At the initial stage, we performed pass-fail prediction by discriminant analysis using intramural mock examination and national examination results for the national radiological technologist examination over a period of one year at a certain school. Specifically, we carried out experiments to investigate the optimal attributes and techniques for prediction by discriminant analysis using the Mahalanobis distance. It was found that the correct identification rate was higher when using the scores obtained in 14 subjects than when using the total of the scores of a mock examination as an attribute. Moreover, the method of dividing the national examination results into two or more groups, conducting discriminant analysis, and performing pass-fail prediction based on those results was determined to be effective. The predictive accuracy of discriminant analysis performed directly on two pass-or-fail groups was 84.3%. On the other hand, the accuracy of pass-fail prediction based on the results of discriminant analysis of five groups was 89.3%, providing a higher predictive accuracy than the former method.

Keywords: Pass-fail prediction, discriminant analysis, national examination, radiological technologist

1. Introduction

National qualifications are required in order for health professionals to perform their work in Japan. It is important for medical schools to have all students pass a national certification examination. In reality, however, not every student passes the examination. There are various reasons for this. To improve this situation, the schools need to provide effective education while assisting individual students to design an appropriate learning program and maintaining their motivation. We focused on educational guidance based on pass-fail predictions for a national examination. This guidance approach allows students, based on pass-fail predictions, to increase awareness of their goals and design an appropriate learning program at the early stages of preparation for the examination. It also allows them to make preparations for the examination while modifying the learning program based on pass-fail predictions conducted on a regular basis. Currently, pass-fail predictions for national examinations are often carried out intuitively based on the experience of faculty members. Data-based objective predictions are considered to be useful in these circumstances.

In the United States, there is a large-scale examination called HESI Exit Exam (E^2) which has been administered by over 600 schools in 50 states (Langford & Willson 2012; Langford & Young, 2013). The study indicated that the accuracy of E^2 in predicting National Council Licensure Examination for Registered Nurses (NCLEX-REN) success is very high (96.36%-99.16%) and the E^2 scores are useful as a benchmark. Unfortunately, there is not such examination for most medical technologist qualification in Japan.

The objective of this research is to develop a method for making pass-fail predictions for a national examination from the results of intramural mock examinations, with utilization by a single school in mind. The national radiological technologist examination was used as an example in this study. As the first step in the development of the method, we evaluated pass-fail predictions by discriminant analysis using intramural mock examination and national examination data collected over a period of one year. Specifically, we studied the attributes to be used and the optimal method of making predictions.

In a study related to pass-fail predictions for national examinations for health care professionals, Richard et al (2004) analyzed the background and scholastic data of undergraduate nursing student who had graduated from a nursing school. They found that gender, age, Scholastic Aptitude Test (SAT) verbal, SAT quantitative and Cumulative Nursing GPA were significantly different between students who had passed NCLEX-REN and students who had not. They created a discriminant function to predict success on NCLEX-REN and the overall hit rate for this function was 70.1%. Truman (2012) compared graduates of a specific rural associate degree nursing program who had been successful on the initial attempt at NCLEX-REN to graduates who had not. As results, it was revealed that Scholastic Aptitude Test (SAT) verbal and nursing GPA significantly predicted NCLEX-RN success.

Miyashita et al (2004) compared average scores in examinations for enrolment in a higher year of study, average grades for 24 subjects in a course-specific graduation examination (course graduation examination), scores in mock examinations for a national examination (prep-school examinations), and scores in a comprehensive final-year examination for final sixth-year students (comprehensive examination) with scores for the national medical practitioners examination, and showed that there is a strong correlation between scores in the prep-school and comprehensive examinations and those in the national examination, indicating that the mock examinations and comprehensive examinations are suitable for predicting scores in the national examination. The previous studies suggest that pass-fail predictions for a national examination can be made more accurately from scores in mock examinations than from those in subject examinations.

Miyamoto et al. (2008) made pass-fail predictions for a national examination from scores in graduation examinations conducted by schools of medicine. The correct pass-fail identification rate for each year ranged from 82.7 to 97.2%. The correct pass-fail identification rate based on the overall average was 91.2%. The rate predicted for the following year using the discriminant for the year of interest ranged from 82.7 to 92.0%, and the overall average was 88.1%. While Miyamoto et al. made predictions from scores in the graduation examination; we evaluated pass-fail predictions, which are used for guidance provided about six months prior to graduation. In addition, we improved the accuracy of the pass-fail prediction by dividing scores in national examination into several groups.

2. Target National Examination and Data Used, and Outline of the Experiments

2.1 Target National Examination

The national examination used in this study was the national radiological technologist examination, described below.

- Held once a year, in February
- Number of questions: 200
- Pass criteria: 60% (120 points) or more and not more than one subject with a score of zero
- Question type: Marked sheet examination in which students select one correct answer from five choices
- 14 examination subjects: General basic medicine, radiation biology (including radiation hygiene), radiological physics, radiochemistry, medical engineering, diagnostic imaging equipment, X-ray photography, diagnostic imaging examination, image engineering, medical imaging informatics, radiation measurement, nuclear medicine examination techniques, radiation therapy techniques, and radiation safety management

2.2 Data Used

The data used in this study were the results of seven intramural mock examinations (including the graduation examination) and the national examination taken by the students graduating from a school in 2011. The data were processed in such a manner as to prevent information from being identified with a particular school and students. We obtained permission for the use of the data for research purposes under the following conditions:

- The data must be modified (student names to be removed and student numbers converted to IDs that cannot be identified with a particular student).
- Information on the source of the data (school name, information provider, number and percentage of students who passed the national examination) must not be disclosed in the results of the study.

An intramural mock examination is prepared by the faculty members. The question type and the number of questions in each subject are the same as those in the national examination. The maximum score is 200 points for both examinations.

A score in each subject is recorded for the mock examination. The total score obtained by post-examination self-assessment is recorded for the national examination. For the national examination, pass and fail information is released by the Ministry of Health, Labor and Welfare, but scores are not released. Therefore, the self-assessed total scores were used for analysis.

2.3 Outline of the Experiments

Pass-fail predictions for the national examination were made by Mahalanobis distance-based discriminant analysis. In a future study, we will obtain information on the probability of individual students passing the national examination and information that can be used to instruct students on which subjects they need to focus on. We will also make pass-fail predictions from the most current results every time a mock examination is conducted. To achieve this, we will make it possible to conduct predictions for each mock examination, instead of making predictions after the results of seven mock examinations become available. The experiments and their objectives are described below.

- Experiment 1: Discriminant analysis of the results of each mock examination. In this experiment, the attributes to be used were determined by comparing two methods for the analysis of mock examination results: a method using a score for each of the 14 subjects and a method using the total score. Moreover, trends in pass-fail predictions from the results of seven mock examinations were identified.
- Experiment 2: Discriminant analysis of the results of all seven mock examinations. The aim of this experiment was to determine the usefulness of pass-fail predictions from the results of all mock examinations as a means of pass-fail prediction for each mock examination.
- Experiment 3: Pass-fail predictions made by dividing national examination results into several groups. In this experiment, it was determined whether national examination results can be predicted at a detailed level. It was also determined whether the accuracy of pass-fail prediction can be improved by making pass-fail predictions from detailed predictions of examination results.
- Experiment 4: Pass-fail predictions without the use of mock examination data in the same session. Mock examination data in the same session as for the prediction sample cannot be used to make pass-fail predictions. Therefore, a similar experiment to Experiment 3 was performed without the use of mock examination data in the same session.

The free statistics software “R” (version R3.0.0) was used for statistical processing. In the discriminant analysis, the equality of the variance-covariance matrix between the pass and fail groups was tested at a significance level of 5%. In the case of equal variance, the Mahalanobis distance was calculated by pooling the variance.

The methods and results of Experiments 1 to 4 are described and discussed in the following sections.

3. Experiment 1: Discriminant Analysis of the Results of Each Mock Examination

First, we focused on the score for each of the 14 subjects and the total score (the sum of the scores for the 14 subjects) to determine the attributes of the intramural mock examination results to be used to effectively determine a pass or fail in the national examination. We performed discriminant analysis with different combinations of these attributes and determined the combination of attributes that gave a high correct identification rate.

3.1 Method

3.1.1 Data Preparation

Data on the results of each examination were prepared as shown in Table 1. The pass-fail attribute is A for a score of 120 or higher (pass) and B for a score of 119 or lower (fail).

Table 1: Example of a data set for Experiment 1

Sample No.	Subject 1	…	Subject 14	Total score in mock exam	Total score in national exam	Pass-fail (group)
1	6	…	11	116	135	A
2	3	…	7	95	119	B
3	5	…	8	125	120	A
4	5	…	10	119	95	B
⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	8	…	10	138	141	A

3.1.2 Testing of Equal Variance

A method for testing the equality of the variance-covariance matrix with the statistics software “R” was applied to obtain the p value. The p values for the seven mock examinations ranged from 0.107 to 0.947. The variance-covariance matrix was determined to be equal between Groups A and B. The variance was pooled to calculate the Mahalanobis distance.

3.1.3 Procedure

The following steps were performed on the data for the first to seventh mock examinations using the score for each of the 14 subjects (14 attributes) and the total score (one attribute).

The Mahalanobis distance from the average for Group A and from the average for Group B was calculated for each sample in Table 1. Predictions were made using the data for the group with a shorter distance as the data set. The results of the predictions were compared with the actual data set for the group, and the number of correct predictions was counted. The correct identification rate was obtained by dividing the number of correct predictions by the total number of data sets.

Table 2: Results of discriminant analysis for each mock examination

Mock Exam / Attribute	1st exam	2nd exam	3rd exam	4th exam	5th exam	6th exam	7th exam
14 Subjects	88.9	84.1	83.6	86.3	88.0	94.7	87.7
Total Score	63.9	71.0	64.2	67.1	57.3	69.3	66.2

3.2 Results

Table 2 shows the correct identification rate for each mock examination obtained using the score for each of the 14 subjects and the total score. As shown in this table, for all mock examinations, the

correct identification rate was higher when the score for each of the 14 subjects was used than when the total score was used.

The correct identification rate obtained using the 14 subjects as attributes gradually decreased for the first to third mock examinations, increased to a maximum of 94.7% for the sixth mock examination, and decreased for the seventh mock examination. In other words, the results show that the accuracy of the prediction is not higher for mock examinations closer to the date of the national examination.

3.3 Discussion

The results of Experiment 1 show that the score for each of the 14 subjects should be used as an attribute for discriminant analysis instead of the total score in a mock examination. The use of the score for each of the 14 subjects may make it possible to analyze which subjects each student should focus on.

From the results of Experiment 1, it is not clear whether the timing of the mock examinations affects the accuracy of the pass-fail prediction. We hypothesized that the accuracy of the pass-fail prediction would be higher for mock examinations closer to the date of the national examination because we considered that the competence of the students would be at a level where they were ready for the national examination. From Table 2, however, it can be seen that the accuracy of the prediction is not necessarily higher for mock examinations closer to the date of the national examination than for earlier mock examinations. This indicates that the accuracy may be more strongly affected by differences in the level of difficulty of the questions and similarity of the content of questions in the national and mock examinations than by proximity to the date of the national examination. There is a weak correlation between the number of examinations and the accuracy of predictions made using the score for each of the 14 subjects (correlation coefficient = 0.46). The accuracy of the prediction may also be affected by information on the date of the national examination.

4. Experiment 2: Discriminant Analysis of the Results of All Seven Mock Examinations

Based on the results of Experiment 1, the score for each of the 14 subjects, which resulted in a high correct identification rate, was used as an attribute. In Experiment 2, the results of all seven mock examinations were used for discriminant analysis. Two cases were analyzed: a case where the scores for each of the 14 subjects were used as attributes and a case where the number of days from the date of the mock examination to the date of the national examination, which comprises information indicating the timing of the mock examination, was added as an attribute (15 attributes in total).

Table 3: Example of a Data Set for Experiment 2

Mock exam	Sample No.	Subject 1	…	Subject 14	No. of days to national exam	Total score in national exam	Pass-fail (group)
1st exam	1	6	…	11	210	135	A
	2	3	…	7	210	119	B
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	n_1	7	…	12	210	141	A
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
7th exam	n_6+1	5	…	10	30	135	A
	n_6+2	4	…	8	30	119	B
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	n_7	8	…	10	30	141	A

The number of the last sample before the m^{th} mock examination is represented by n_m.

4.1 Method

4.1.1 Data Preparation

Data on the results of all seven examinations were prepared as shown in Table 3. The number of days to the date of the national examination was added as an attribute. The number of days to the date of the national examination indicates how many days before the national examination the mock examination was conducted. The pass-fail attribute is A for a total score of 120 or higher (pass) and B for a total score of 119 or lower (fail) in the national examination.

4.1.2 Testing of Equal Variance

The method for testing the equality of the variance-covariance matrix with the statistics software "R" was applied to obtain the p value. As a result, the p value was 0.0129 for the score for each of the 14 subjects (14 attributes without the number of days to the national examination) and 0.0196 for those with the number of days included (15 attributes). The variance-covariance matrix was determined to be not equal between Groups A and B. Therefore, the variance was not pooled to calculate the Mahalanobis distance.

4.1.3 Procedure

As with Experiment 1, the correct identification rate was obtained using the data on all seven mock examinations in two cases: a case where the score for each of the 14 subjects was used (case without the use of the number of days) and a case where the score for each of the 14 subjects and the number of days to the date of the national examination were used (case with the use of the number of days).

4.2 Results

The correct identification rate was 88.1% in the case without the use of the number of days and 89.3% in the case with the use of the number of days. The accuracy of the prediction was slightly higher in the latter case.

4.3 Discussion

It was found that if the data for the results of all mock examinations were processed in a batch, the variance-covariance matrix of the mock examination results was not equal between the pass and fail groups. Therefore, the variance was not pooled to calculate the Mahalanobis distance.

The accuracy of the prediction was slightly higher in the case with the use of the number of days to the date of the national examination, which comprises information indicating the timing of a mock examination. However, it is not clear whether this information contributes to the accuracy of the pass-fail prediction due to the proximity to the date of the national examination. In this experiment, to make pass-fail predictions for a sample, mock examination data in the same session as for the sample were included in the data set. As a result, the effect of the mock examination data in the same session on the accuracy of the prediction increased due to the use of the number of days, and the possibility of correct predictions increased.

5. Experiment 3: Pass-fail Predictions Made by Dividing National Examination Results into Several Groups

In Experiments 1 and 2, discriminant analysis was performed on the pass and fail groups to determine the accuracy of the prediction. In Experiment 3, the number of national examination score groups was increased to make more detailed predictions of national examination results, and discrimination analysis (multiple discriminant analysis) was performed. The accuracy of pass-fail (two groups) predictions from the results of the analysis was evaluated.

5.1 Method

5.1.1 Grouping

Table 4 shows the grouping. The scores in the national examination were divided according to the number of groups, as shown in the table. That is, students were grouped by the range of their scores in the national examination shown in Table 4. For example, in the case of four groups, a score from 140 to 200 inclusive falls within Group A and a score of 130 to 139 inclusive falls within Group B.

In the case of two groups, students were divided into only two groups: passing students and failing students. In the case of three and four groups, passing students were divided into several groups. In the case of five groups, failing students were divided. The students were divided into groups so that the number of samples was not zero in any case.

Table 4: Range of the Score of Each Group

No. of Groups	Group A	Group B	Group C	Group D	Group E
2	200~120	119~0	-	-	-
3	200~135	134~120	119~0	-	-
4	200~140	139~130	129~120	119~0	-
5	200~140	139~130	129~120	119~110	109~0

Note: The painted score ranges indicate a pass.

5.1.2 Data Preparation

Data for the results of all seven examinations were prepared as shown in Table 5. The pass-fail attributes were grouped in the same manner as in Table 4. Table 5 shows an example of a data set in the case of five groups.

Table 5: Example of a data set for Experiment 3

Mock exam	Sample No.	Subject 1	…	Subject 14	Total score in national exam	Group
1st exam	1	6	…	11	89	E
	2	3	…	7	119	D
	⋮	⋮	⋮	⋮	⋮	⋮
	n_1	7	…	12	141	A
⋮	⋮	⋮	⋮	⋮	⋮	⋮
7th exam	n_6+1	5	…	10	89	E
	n_6+2	4	…	8	119	D
	⋮	⋮	⋮	⋮	⋮	⋮
	n_7	8	…	10	141	A

5.1.3 Procedure

As with Experiment 1, the correct identification rate was obtained for a different number of national examination score groups from 2 to 5 in two cases: the case where the score for each of the 14 subjects was used (case without the use of the number of days) and the case where the score for each of the 14 subjects and the number of days to the date of the national examination were used (case with the use of the number of days). In addition, pass-fail predictions were made for each group using Table 4. The predictions were compared with the actual pass-fail results, and the number of correct predictions was counted to determine the accuracy of the prediction.

5.2 Results

Table 6 shows the correct group identification rate and the correct pass-fail prediction rate for different numbers of national examination score groups. The group correct identification rate is the highest when the number of groups is 2 and decreases with an increasing number of groups. The

correct pass-fail prediction rate is higher for a larger number of groups and is the highest when the number of groups is 5. For all numbers of groups, the rate is slightly higher in the case with the use of the number of days.

Table 6: Results of Group Distinction and Pass-Fail Prediction

No. of Groups	Without No. of Days to National Exam		With No. of Days to National Exam	
	Group Distinction	Pass-Fail Prediction	Group Distinction	Pass-Fail Prediction
2	88.1	88.1	89.3	89.3
3	68.1	89.7	69.8	91.1
4	61.1	90.9	66.1	92.7
5	64.3	96.0	68.5	96.8

5.3 Discussion

The results of Experiment 3 show that if the number of national examination score groups is more than 2, the correct prediction rate based on score groups decreases. Therefore, it is difficult to use this method to make detailed predictions of examination results. Pass-fail predictions made based on the predicted groups are more accurate than those made based directly on the pass and fail groups. Particularly when the number of groups is 5, the correct prediction rate is at least 7 points higher than when the number of groups is 2. This indicates that pass-fail predictions based on the grouping of national examination results are useful.

As in Experiment 2, the accuracy of the prediction was slightly higher in the case with the use of the number of days to the date of the national examination in Experiment 3. However, as discussed in Section 4.3, whether to use the number of days to the date of the national examination as an attribute is an issue that needs to be examined.

6. Experiment 4: Pass-fail Predictions Without the Use of Mock Examination Data in the Same Session

Mock examination data in the same session as for the prediction sample cannot be used to make pass-fail predictions for a forthcoming national examination. The experiments conducted indicate (a) that the accuracy of pass-fail prediction increases due to the effect of the mock examination data in the same session and (b) that the accuracy of the prediction slightly improves because the effect of the data increases when the number of days to the date of the national examination is used. Therefore, a similar experiment to Experiment 3 was performed without the use of mock examination data in the same session as for the prediction sample.

6.1 Method

The grouping and data preparation were the same as in Experiment 3.

The following steps were performed for different numbers of national examination score groups from 2 to 5 in two cases: the case where the score for each of the 14 subjects was used (case without the use of the number of days) and the case where the score for each of the 14 subjects and the number of days to the date of the national examination were used (case with the use of the number of days).

The data in Table 5, excluding the first examination data, were used for discrimination. The Mahalanobis distance from the average for each group was calculated using discrimination data for each sample from the first examination. Predictions were made using the data for the group with the shortest distance as the data set. The results of the predictions were compared with the actual data for the group, and the number of correct predictions was counted for correct group discrimination. Pass-fail predictions were made based on whether the predicted group was a pass group or a fail group, and were compared with the actual pass-fail results. The number of correct predictions was counted for correct pass-fail predictions. The above steps were performed using the first to seventh

data sets. The correct group identification rate and the correct pass-fail prediction rate were obtained by dividing the number of correct group identifications and the number of correct pass-fail predictions by the total number of samples.

Table 7: Results of Group Distinction and Pass-Fail Prediction Without Mock Examination Data to Which the Target Sample Belonged

No. of Groups	Without No. of Days to National Exam		With No. of Days to National Exam	
	Group Distinction	Pass-Fail Prediction	Group Distinction	Pass-Fail Prediction
2	82.7	82.7	84.3	84.3
3	50.0	82.9	55.4	84.9
4	37.5	82.7	41.1	83.5
5	39.1	88.5	43.1	89.3

6.2 Results

Table 7 shows the results. Overall, the accuracy of the prediction was lower than in Experiment 3, in which all data were used. Unlike in Experiment 3, the correct pass-fail prediction rate did not necessarily increase with an increasing number of groups, but was the highest when the number of groups was 5. The accuracy of the prediction for all groups was slightly higher in the case with the use of the number of days.

6.3 Discussion

The correct prediction rate is about 5 points lower in the case without the use of mock examination data in the same session as that for the prediction sample, compared with the case in which the entire mock examination data are used. This indicates that the accuracy of the prediction will be 80 to 90% using discriminant analysis to make pass-fail predictions for the national examination. Since the level of difficulty of the questions in the national examination differs from year to year, we will continue to evaluate pass-fail predictions using data for several years.

The results of Experiment 4 indicate that the accuracy of the prediction slightly increases using the number of days from the date of the mock examination to the date of the national examination. This is because in Experiment 4, mock examination data in the same session as for the prediction sample were not used and the increased effect of the mock examination data in the same session on the accuracy of the prediction due to the use of the number of days did not occur.

7. Conclusions

In the initial stage of this study of pass-fail predictions for the national radiological technologist examination, we performed Mahalanobis distance-based discriminant analysis using data on examination results collected over a period of one year. Our analysis indicates that predictions can be made with an accuracy approaching 90% and has shown the following:

- The score for each of the 14 subjects in a mock examination should be used as a basic attribute.
- Pass-fail predictions can be effectively made by dividing national examination results into five groups.
- The number of days to the date of the national examination may be a useful attribute for pass-fail predictions.

Data collected over a period of two years will be available for a future study. We will evaluate the accuracy of the prediction by making pass-fail predictions for the following year using data for the current year. The next step is to make pass-fail predictions using other techniques such as multiple regression analysis, neural net analysis, and Bayesian analysis.

References

Langford,R. & Willson,P. (2012) Predicting NCLEX-RN Success: The Seventh Validity Study of the HESI Exit Exam, *CIN: Computers, Informatics, Nursing, 30*, 55-60.

Langford,R. & Young,A. (2013) Predicting NCLEX-RN Success With the HESI Exit Exam: Eighth Validity Study, *Journal of Professional Nursing, 29*(25), 55-59.

Miyamoto, M., Mori, Y., Kubota, T., Nishimura, Y., & Yoneda, H. (2008). Expectation of Graduates' Pass-fail Possibilities on the National Examination by Discriminant Analysis of the Results on the Graduation and Integrated Examinations, *The Japanese Journal for Research on Testing, 4*(1), 34-39.

Miyashita, T., Shimura, T., Adachi, K., Aramaki, T., Shimizu, K., & Dan, K.(2004). A Comparative Study of Undergraduate Examinations to Evaluate Medical Students' Performance and the National Examination for Medical Practitioners. *Journal of the Japanese Society for Medical Education, 35*(4), 281-285.

Haas,R.E., Nugent,K.E., & Rule R.A. (2004) The Use of Discriminant Function Analysis to Predict Student Success on the NCLEX-REN, *Journal of Nursing Education, 43*(10), 440-446.

Truman,J.M. (2012) Identifying Predictors of National Council Licensure Examination for Registered Nurses (NCLEX-RN) Success in an Associate Degree Nursing Program, *International Journal of Applied Science and Technology, 2(7)*,37-45.