CREATIVE MATH.
14 (2005), 67 - 70

The estimation of the accuracy of an indicator in educational testing theory, using the bootstrap method

Cristian Marinoiu

Abstract. Often, the researcher in educational testing theory, would like to have a model that simply explains the obtained scores. In order to establish if such a model is adequate, an indicator based on the eigenvalues of the empirical covariance matrix of the scores is used. This paper presents the standard methodology to construct this indicator, its bootstrap estimate of the standard error and its percentile bootstrap confidence interval. Finally, for a case study, a program that implements in Matlab the described algorithm provides the desired results.

1. Introduction

Let us consider that each of m random selected students take n tests in different subjects. For easiness, we organize the obtained scores data in a $X(m \times n)$ matrix, when the x_{ij} component represents the score obtained by the student i in the subject j.

A good measure of the degree of dependence between the scores is the empirical correlation matrix, defined as

$$C_{jk} = \frac{\frac{1}{m}\sum\limits_{i=1}^{m}(x_{ij}-\bar{x}_j)\cdot(x_{ik}-\bar{x}_k)}{\left(\sum\limits_{i=1}^{m}(x_{ij}-\bar{x}_j)\right)^{\frac{1}{2}}\cdot\left(\sum\limits_{i=1}^{m}x_{ik}-\bar{x}_k\right)^{\frac{1}{2}}} \quad j,\, k = 1,\, 2,\, ...,\, n \tag{1}$$

where

$$\bar{x}_j = \left(\sum_{i=1}^{m} x_{ij}\right)/m$$

If, after studying the empirical correlation matrix we conclude that there is a high dependence between the n vectors of the scores, we will encouraged to think a simple explanatory model for the correlated scores. Such a model would be, see for example [1],

$$X_i = K_i v \quad i = 1,\, 2,\, ...,\, m \tag{2}$$

where:

$X_i = (x_{i1}, x_{i2}, ..., x_{in})$ is the i^{th} row of the matrix X,

K_i is a scalar representing the capability of the student i,

v is a fixed vector, applying to all students.

Received: 10.05.2004. In revised form: 15.05.2005.
2000 *Mathematics Subject Classification.* 62F25.
Key words and phrases. *Educational testing theory, eigenvalues, bootstrap method.*

A measure of the degree of confidence in such a model is provided by the eigenvalues of the empirical covariance matrix of the scores data G, defined as:

$$G_{jk} = \frac{1}{m} \sum_{i=1}^{m} (x_{ij} - \bar{x}_j) \cdot (x_{ik} - \bar{x}_k) \qquad j,\, k = 1,\, 2,\, ...,\, n \tag{3}$$

Let us denote by $\hat{\lambda}_1,\, \hat{\lambda}_2,\, ...,\, \hat{\lambda}_n$ these eigenvalues with $\hat{\lambda}_1 \geqslant \hat{\lambda}_2 \geqslant\, ... \geqslant \hat{\lambda}_n$ and by $\hat{v}_1,\, \hat{v}_2,\, ...,\, \hat{v}_n$ the corresponding eigenvectors. The eigenvector $\hat{v}_1$ is called first principal component of the matrix G.

An equivalent form of the model (2) in terms of the eigenvalues is

$$\hat{\lambda}_1 > 0,\, \hat{\lambda}_2 = \hat{\lambda}_3 = \, ... \, = \hat{\lambda}_n = 0\,, \;\; v = \hat{v}_1,$$

or

$$\hat{r} = \hat{\lambda}_1 / \left(\sum_{i=1}^{n} \hat{\lambda}_i \right) = 1$$

In reality, $\hat{r} \in (0,\, 1]$, and therefore it can be considered as an indicator measuring the distance between the model and the real situation provided by data. Obviously, the closer $\hat{r}$ is to 1, the higher our confidence in the model (2).

In order to establish how accurate $\hat{r}$ is, since it is not a neat formula to estimate its standard error, a convenient way is to use the bootstrap method.

Moreover, the bootstrap method allows us to obtain a confidence interval, for the true value of r, which is estimated by $\hat{r}$, namely

$$r = \lambda_1 / \left(\sum_{i=1}^{n} \lambda_i \right)$$

where $\lambda_i, i = 1, 2, ..., n$ are the eigenvalues of the unknown covariance matrix of the random variables for which we have the observed scores.

2. The bootstrap estimate of standard error and the percentile bootstrap interval

Consider we know a random sample $x = (x_1,\, x_2,\, ...,\, x_n)$ from an unknown probability distribution F.

If

$$\widehat{F} = \begin{pmatrix} x_1 & x_2 & ... & x_n \\ \dfrac{1}{n} & \dfrac{1}{n} & ... & \dfrac{1}{n} \end{pmatrix}$$

is the empirical distribution function one can estimate some interesting aspects of F using $\hat{F}$ (the plug-in principle). Let us notice that, while from the unknown probability distribution F is available only the sample $x = (x_1,\, x_2,\, ...,\, x_n)$, from $\hat{F}$ one can obtain C_{2n-1}^{n} distinct bootstrap samples, $x^* = (x_1^*, x_2^*,\, ...,\, x_n^*)$.

Each component of the bootstrap sample x^* is obtained by drawing with replacement from the components of the random sample $(x_1,\, x_2,\, ...,\, x_n)$. Let us denote by θ the parameter of interest, and let be:

$\hat{\theta} = s(x)$ an estimation for θ

$\hat{\theta}^* = s(x^*)$ the value of the statistics $s(t)$, for $t = x^*$.

Thus, an estimate of the standard error of the statistics of interest $\hat{\theta}$, is:

$$\widehat{se}_B = \left[\sum_{b=1}^{B} \left(\hat{\theta}^* (b) - \hat{\theta} (\cdot) \right)^2 / (B-1) \right]^{\frac{1}{2}} \tag{4}$$

where:

B is the number of independent bootstrap samples

$\hat{\theta}^*$ is the b^{th} bootstrap replication of $\hat{\theta}$

$$\hat{\theta} (\cdot) = \sum_{b=1}^{B} \hat{\theta}^* (b) / B.$$

The $1-2\alpha$ percentile interval, see [1], is defined by the α and $1-\alpha$ percentiles of $\hat{F}^*$, as

$$\left[\hat{F}^{*-1} (\alpha) \right] \ , \ \left[\hat{F}^{*-1} (1-\alpha) \right] \tag{5}$$

where $\hat{F}^*$ is the cumulative distribution function of $\hat{\theta}^*$.

3. The bootstrap algorithm

Let us notice that, in our case, the parameter of interest is r and the estimator is $\hat{r}$. Therefore the steps to obtain the desired results are as follows:

1. Read $\alpha, m, n, X(m \times n)$

2. Generate B bootstrap independent samples $\left(X_1^{*1}, X_2^{*2}, ..., X_n^{*B} \right)$ from the rows

$$X_1, X_2, ..., X_m \text{ of the matrix } X.$$

3. For each bootstrap sample:

Form the bootstrap matrix X^*

Form the empirical covariance matrix G_b^* using the formula (3)

Calculate the eigenvalues of G_b^*, as $\hat{\lambda}_1^*, \hat{\lambda}_2^*, ..., \hat{\lambda}_n^*$, where

$$\hat{\lambda}_1^* \geqslant \hat{\lambda}_2^* \geqslant ... \geqslant \hat{\lambda}_n^*$$

Calculate the bootstrap replications of $\hat{r}$,

$$\hat{r}^* = \hat{\lambda}_1^* / \left(\sum_{i=1}^{n} \hat{\lambda}_i^* \right)$$

4. Calculate the estimate of the standard error of $\hat{r}$ and a percentile interval for r using the formulae (4) and (5) respectively

5. Stop.

4. Results and conclusions

Our study refers to the scores obtained by $m = 19$ students in the winter session exams, in $n = 3$ subjects: Fundamentals of Informatics (FI), Algebra (A) and Logic and Theory of sets (LT).

Table 1. The score data

No.	FI	A	LT
1.	8	4	9
2.	7	4	9
3.	9	4	8
4.	7	5	8
5.	10	6	9
6.	8	5	9
7.	4	4	2
8.	7	4	8
9.	6	5	9
10.	8	7	8
11.	10	5	10
12.	7	4	5
13.	4	4	7
14.	4	4	8
15.	10	4	9
16.	8	5	7
17.	8	4	9
18.	7	4	8
19.	6	4	7

The scores in Table 1, serve as entrance data for the Matlab program *bootted* that implements the bootstrap algorithm. We have chosen the Matlab software because is convenient in operations with matrices (the calculation of the eigenvalues, for example).

For $B = 200,\ \alpha = 0.05,$ the obtained results are as follows:

$\hat{r} = 0.7313$

the bootstrap standard estimate of $\hat{r}$ is 0.07645

the percentile bootstrap confidence interval is $(0.6019,\ 0.8448)$

The obtained value for $\hat{r}$ and its high accuracy indicate a good degree of explanatory power for model (2).

References

[1] Bradley E., Tibshirani R., *An introduction to the bootstrap*, Chapman & Hall, USA, 1993

[2] Demidovich B.P., Maron L.A., *Computational mathematics* , Mir Publishers, Moscow, 1981

University of Ploieşti
Petroleum - Gas University of Ploieşti
Bd. Bucureşti 39, 100680 Ploieşti, Romania
E-mail address: marinoiu_c@yahoo.com