Research on the Difference Significance of Test Score by SAS

Huameng Gao\textsuperscript{a}, Yong Xue\textsuperscript{*}, Yu Sun\textsuperscript{b}, Jing Wang\textsuperscript{c}

Space Engineering University, Beijing, China

\textsuperscript{*}Corresponding author e-mail: xueyong0262@163.com, gao0539gao0539@126.com, sunyu930821@sina.cn, 13366023699@189.cn

Abstract. In order to research the significancy level and difference in different groups of test score data, the way for checking data’s normal character is introduced. In order to realizing variance analysis, t test and Wilcoxon test, the SAS function, such as SORT, UNIVARIATE NORMAL, TTEST COCHRAN and NPAR1WAY WILCOXON are presented. According to real data, the distribution histgraph, $X^2$ values, normal test, signed-rank test, and Wilcoxon two-sample test are gained. The study in this paper can be adopted in other fields of test design and data analysis.

1. Introduction
Evaluating the difference significance of test score in different groups is a key problem in education and is the basis for examination system. The checking and processing to test score have been researched in many papers. For example, the $X^2$ checkout for test score is researched by Liu Xiaohui [1]. The numerical analysis for test score through Exel is researched by Liu Donghua and Pan Qufeng [2]. The distribution and it’s causation of student’s test score are researched by Sun Weiquan et al [3]. The transform and synthesizing of different test scores are researched by Liu Xuezhi [4]. The distribution and exporting of test score are researched by He Minggao and Wu Fugen [5]. The numerical evaluating way to the normal of test score is researched by Gao Jie [6]. In these papers, the processing way to a block of test score is researched, but the difference significance in different blocks of test score is ignored. In this paper, all of the above two problem are researched, and the content of this paper can be used in other fields of data analysis.

2. The problem
There are two classes all with 100 students, and the two groups of students’ test score are as below:

Class 1: 81 68 86 88 74 97 97 85 88 87 83 92 79 100 84 86 96 86 84 77 88 72 92 100 78 94 98 69 71 91 81 92 93 92 98 92 97 73 85 83 69 88 74 99 77 90 87 76 63 84 75 91 90 100 91 79 89 75 85 85 82 96 66 89 94 92 91 85 92 91 82 81 82 70 83 86 88 99 81 91 93 94 75 87 75 78 96 84 89 86 79 79 89 76 93 91 77 82

Class 2: 72 62 94 79 87 86 84 74 75 80 72 73 99 85 72 84 73 71 81 94 85 91 72 79 81 72 71 93 84 78 92 86 74 97 87 99 95 77 71 83 81 76 89 99 79 76 82 91 75 72 83 83 89 78 83 80 83 99 78 71 89 75 84 78 89 90 82 63 85 99 94 68 83 77 74 72 87 80 85 80 79 88 91 99 70 74 94 80 70 87 99 91 99 89 99 81 73 82 96 73

Now, the difference significanse of these two groups data should be analized and the better class should be found.
3. The analysis procedure

3.1. normal check
The hist of Class 1 is as figure 1:

![Figure 1. The histgraph of Class 1.](image)

In the first group data, the maximum is 100, and the minus is 63. In figure 1, the axes X presents score, and the axes Y presents frequency. The scores are divided to 10 groups with 3.7 distance.

The average and standard variance are as below:

\[ X = \frac{\sum_{i=1}^{n} X_i}{n} \]  \hspace{1cm} (1)

\[ S = \left( \frac{\sum_{i=1}^{n} (X_i - X)^2}{n} \right)^{0.5} \]  \hspace{1cm} (2)

For class 1, the \( X \) is 85.4, and the \( S \) is 8.4427.

The probability in every dividing number can be gained by \( F(X_i) = \phi\left(\frac{X_i - X}{S}\right) \), \( \phi \) is the distribution function of normal distribution.

\[ F(63) = 0.0040 \quad F(66.7) = 0.0134 \quad F(70.4) = 0.0378 \quad F(74.1) = 0.0904 \]  \hspace{1cm} (3)

\[ F(77.8) = 0.1840 \quad F(81.5) = 0.3221 \quad F(85.2) = 0.4906 \quad F(88.9) = 0.6602 \]  \hspace{1cm} (4)

\[ F(92.6) = 0.8031 \quad F(96.3) = 0.9017 \quad F(100) = 0.9581 \]  \hspace{1cm} (5)

The \( X^2 \) values of the data in the first group is as table 1.
Table 1. The $X^2$ values of the data in the second group

| block          | frequency(v) | P     | nP    | $(v-np)^2$ | $(v-np)^2/np$ |
|----------------|-------------|-------|-------|------------|--------------|
| [63.0 66.7]    | 2           | 0.0094| 0.94  | 1.1236     | 1.1953       |
| [66.7 70.4]    | 4           | 0.0244| 2.44  | 2.4336     | 0.9974       |
| [70.4 74.1]    | 5           | 0.0526| 5.26  | 0.0676     | 0.0129       |
| [74.1 77.8]    | 9           | 0.0936| 9.36  | 0.1296     | 0.0138       |
| [77.8 81.5]    | 10          | 0.1381| 13.81 | 14.5161    | 1.0511       |
| [81.5 85.2]    | 17          | 0.1685| 16.85 | 0.0225     | 0.0013       |
| [85.2 88.9]    | 14          | 0.1696| 16.96 | 8.7616     | 0.5166       |
| [88.9 92.6]    | 20          | 0.1429| 14.29 | 32.6041    | 2.2816       |
| [92.6 96.3]    | 9           | 0.0986| 9.86  | 0.7396     | 0.0750       |
| [96.3 100]     | 10          | 0.0564| 5.64  | 19.0096    | 3.3705       |
| total          | 100         |       |       | 9.5155     |              |

Because $9.5155 < 14.067$, so the data in the first group are accord to normal distribution.

The significance level $\alpha = 0.005$ is selected, and the df is 7. The value of $X^2_{0.05,7} = 14.067$. Because $9.5155 < 14.067$, so the data in the first group are accord to normal distribution.

The hist of the second group of data is as figure 2.

![Figure 2](image)

Table 2. The $X^2$ values of the data in the second group

| block          | frequency(v) | P     | nP    | $(v-np)^2$ | $(v-np)^2/np$ |
|----------------|-------------|-------|-------|------------|--------------|
| [50.0 62.2]    | 2           | 0.0089| 0.89  | 1.2417     | 1.4019       |
| [62.2 66.4]    | 0           | 0.0240| 2.40  | 5.7611     | 2.4002       |
| [66.4 70.6]    | 9           | 0.0532| 5.32  | 13.5188    | 2.5396       |
| [70.6 74.8]    | 9           | 0.0966| 9.66  | 0.4385     | 0.0454       |
| [74.8 79.0]    | 10          | 0.1435| 14.35 | 18.9586    | 1.3208       |
| [79.0 83.2]    | 21          | 0.1745| 17.45 | 12.5756    | 0.7205       |
| [83.2 87.4]    | 17          | 0.1737| 17.37 | 0.1375     | 0.0079       |
| [87.4 91.6]    | 12          | 0.1415| 14.15 | 4.6236     | 0.3267       |
| [91.6 95.8]    | 11          | 0.0943| 9.43  | 2.4507     | 0.2598       |
| [95.8 100]     | 9           | 0.0515| 5.15  | 14.8347    | 2.8814       |
| total          | 100         |       |       | 11.9042    |              |

Because $11.9042 < 14.067$, so the data in the second group are accord to normal distribution.
3.2. **SAS program**

It has difference between the data of the two group, but if the difference is significant and has statistics significant should be tested through statistics way, such as variance analysis, t test and Wilcoxon test.

In this paper, the variance analysis, t test and Wilcoxon test are realized in SAS. The SAS program consists of data content and procedure content. In data content, the data set is constructed. In procedure content, the functions such as SORT, UNIVARIATE NORMAL, TTEST COCHRAN and NPAR1WAY WILCOXON are used. The results of SAS program are as bellow.

The result of the normal test is as table 3.

**Table 3. the result of the normal test**

| test        | statistics parametre | value     | P       | the P value |
|-------------|----------------------|-----------|---------|-------------|
| Shapiro-Wilk| W                    | 0.977189  | Pr < W  | 0.0801      |
| Shapiro-Wilk| W                    | 0.967461  | Pr < W  | 0.0142      |

Through Shapiro-Wilk test, the data of group 1 is accord to normal distribution, but the data of group 2 is not accord to normal distribution. In this condition, through signed-rank test, the difference between the anverages of the data in two groups can be test.

The result of the signed-rank test is as table 4 and 5.

**Table 4. the result of the signed-rank test for data 1**

| test    | statistics parametre | value     | P       | the P value |
|---------|----------------------|-----------|---------|-------------|
| t       | t                    | 100.6449  | Pr > |t|   | <.0001     |
| sign    | M                    | 50        | Pr >= |M|   | <.0001     |
| signed-rank | S            | 2525     | Pr >= |S|   | <.0001     |

**Table 5. the result of the signed-rank test for data 2**

| test    | statistics parametre | value     | P       | the P value |
|---------|----------------------|-----------|---------|-------------|
| t       | t                    | 90.66981  | Pr > |t|   | <.0001     |
| sign    | M                    | 50        | Pr >= |M|   | <.0001     |
| signed-rank | S            | 2525     | Pr >= |S|   | <.0001     |

The Wilcoxon two-sample test result is as bellow.

Statistic: 11043.0000
Normal Approximation
Z: 2.4267
One-Sided Pr > Z 0.0076
Two-Sided Pr > |Z| 0.0152

The Wilcoxon test result is as bellow.

Statistic: 11043.0000
Normal Approximation
Z: 2.4267
One-Sided Pr > Z 0.0076
Two-Sided Pr > |Z| 0.0152

Accord to the result of signed-rank test and Wilcoxon test result, the averages of the two data groups have statistics difference.

4. **Conclusion**

Evaluating the difference significance of test score in different groups is a key problem in education and is the basis for examination system. X² values calculating can be used for checking normal character. In order to realizing more analysis for test score, the t test and Wilcoxon test should be introduced, and the SAS program and SAS function, such as SORT, UNIVARIATE NORMAL, TTEST COCHRAN and NPAR1WAY WILCOXON should be used.
Reference
[1] Liu Xiaohui. The X² Test for Test Score in 150 Degree[J]. Journal of Shenzhen Institute of Education, 2000, (1):19-21.
[2] Liu Donghua, Pan Qiufeng. The Simple Quantitative Analysis for Test Score Based on Excel[J]. China Education Info, 2011.02:57-58.
[3] Sun Weiquan, Deng Deming, Peng Xiandao, et al. A Study on the Distribution and Causation of the Students’ Test Score[J]. Northwest Medical Education, 2000, 8(4):198-200.
[4] Liu Xuezhi. Research on the Mistake in the Exchanging and Composing Between Different Subject’s Test Score[J]. China Examinations, 2010, 39(6):14-18, 22.
[5] He Minggao, Wu Fugen. The Test Score Distribution, Original Test Score and Educated Test Score[J]. Education Exploring, 2000, (3):21-22.
[6] Gao Jie. The Quantitative Evaluation for Test Score Distribution’s Normal Character[J]. Technology Economics, 1998, (8):51-52.