

Bi-Aspect Nonparametric Testing

Test Nonparametrici Bi-Aspetto

Marco Marozzi

Istituto di Statistica Medica e Biometria, Università di Milano
via Venezian 1, 20133 Milano, tel. 02 50320855, e.mail marozzi@stat.unibo.it

Riassunto: facendo riferimento alla teoria della Combinazione Nonparametrica (Pesarin 2001), Marozzi (2002 & 2003) ha proposto due metodi nonparametrici bi-aspetto per il controllo di ipotesi sulla locazione. Esperimenti di simulazione rivelano come questi test siano marcatamente più potenti dei tradizionali test parametrici e di permutazione (tipicamente test uni-aspetto) quando la distribuzione generatrice è a coda pesante o notevolmente asimmetrica. In questo lavoro si fornisce una presentazione dettagliata e unificata della teoria e dell'idea alla base dei metodi. Vengono inoltre discusse alcune applicazioni a dati reali che mostrano come i test bi-aspetto permettano di condurre un'inferenza maggiormente informativa di quella ottenibile con strumenti tradizionali.

Keywords: Nonparametric Methods, Permutation Testing, Bi-Aspect Testing, Nonparametric Combination of Dependent Tests, Applications.

1. Introduction

Starting by considering that traditional parametric and permutation tests for the C -sample location problem ($C \geq 2$) can be seen as uni-aspect tests, Marozzi (2002 & 2003) proposed the permutation bi-aspect tests T_{ab} and T_{AB} , respectively for the two-sample and the multi-sample location problem. Traditional tests take into account only the numerical aspect X_i of each sample unit, whereas bi-aspect tests consider also the aspect related to the comparison of sample units with the median of the pooled sample. Marozzi (2002 & 2003) showed that the bi-aspect tests are remarkably more powerful than traditional permutation tests under heavy-tailed and/or asymmetric distributions like the Cauchy, half-Cauchy, 10% and 30% outlier distribution. When sampling from the double-exponential and exponential distributions, the bi-aspect tests appear to be better on the whole than traditional permutation tests. Under normal, chi-squared, uniform and bimodal they are practically as powerful as traditional tests. Moreover the bi-aspect tests maintain the type-one error rate close to the nominal significance level and they are generally slightly conservative.

The reason for these good results is that the bi-aspect tests are devised in order to take the best from the two partial tests. The first partial test is the traditional permutation test and uses the totals (i.e. the means) to detect location shifts, that is a good choice for example under normal distributions. The second partial test uses frequencies of observations above the pooled median, and for this reason it performs much better than the traditional permutation tests under heavy-tailed and highly-skewed distributions, for which sample means are not very sensitive in studying locations (Johnson and Kotz, 1970 and Huber, 1981).

2. The Bi-Aspect Testing Idea

Let $\underline{X}_1 = (X_{11}, \dots, X_{1n_1})$ and $\underline{X}_2 = (X_{21}, \dots, X_{2n_2})$ be independent random samples from populations with continuous distribution functions that may differ only in their locations. Let $\underline{X}$ denote the pooled sample, $\underline{X} = (X_{11}, \dots, X_{1n_1}, X_{21}, \dots, X_{2n_2})$, $\underline{X}^*$ a permutation of $\underline{X}$, $\underline{X}^* = (X_{11}^*, \dots, X_{1n_1}^*, X_{21}^*, \dots, X_{2n_2}^*)$ and $n = n_1 + n_2$ the total sample size. To test the hypothesis $H_0 : \delta_1 = \delta_2$ against the alternative $H_0 : \delta_1 > \delta_2$, where δ_h denotes the location parameter of population h ($h=1, 2$), Marozzi (2002) proposed the test

$$T_{ab} = \max(1 - L_{T_a}, 1 - L_{T_b}),$$

where L_{T_a} is the p -value of the permutation partial test $T_a = \sum_{i=1}^{n_1} X_{1i}^*$ and L_{T_b} is the p -value of the permutation partial test $T_b = \sum_{i=1}^{n_1} I(X_{1i}^* > \tilde{M})$, where $\tilde{M}$ denotes the median of $\underline{X}$ and $I(\cdot)$ the indicator function. L_{T_a} , L_{T_b} and the p -value $L_{T_{ab}}$ of T_{ab} are permutationally estimated. T_a and T_b are dependent in a way that is generally difficult to assess and their joint distribution, together with that of T_{ab} , is evaluated by a conditional Monte Carlo procedure carried out a number of times, since the set of all data permutations is generally impractically large (see Marozzi 2002, for more technical details and a step by step computing algorithm).

T_{ab} is obtained in the framework of the Nonparametric Combination of Dependent Tests (Pesarin, 2001). T_{ab} is a bi-aspect test because it combines T_a and T_b tests. T_a acts on the first sample total (which is permutationally equivalent to the Student's t statistic) and takes into account the numerical aspect X_{1i}^* . T_b acts on frequencies above the pooled sample median and takes into account the dichotomous aspect $I(X_{1i}^* > \tilde{M})$ related to the fact that a sample unit is small or large with respect to the other ones. An alternative choice of the cut-off point for the second aspect may be the pooled sample mean. It should be noted that we used the median rather than the mean of the pooled sample because the median is intrinsically more robust than the mean. Moreover in our opinion, for what concerns the bi-aspect testing structure it is not superfluous that the median produces a more balanced division of sample elements. Finally, we think that this choice is the most natural one to take into account the dichotomous aspect. Traditional tests use the sample mean to detect location shifts, but in the presence of outliers and heavily skewed data, the mean is not a proper statistic to study location. In these cases, due to its full robustness (Huber, 1981) the median is more sensitive than the mean in detecting location shifts.

Marozzi (2003) proposed a natural extension of T_{ab} to the multi-sample location problem. Let $\underline{X}_1 = (X_{11}, \dots, X_{1n_1})$, $\underline{X}_2 = (X_{21}, \dots, X_{2n_2})$, ..., $\underline{X}_C = (X_{1C}, \dots, X_{1n_C})$ be $C \geq 2$ independent random samples drawn from populations with continuous distribution functions that may differ only in their locations $\delta_1, \delta_2, \dots, \delta_C$, $n = \sum_{j=1}^C n_j$. Let the hypotheses under testing be $H_0: \delta_1 = \delta_2 = \dots = \delta_C$ against $H_1: \delta_h \neq \delta_j$ for some pair (h, j) . The bi-aspect test for the multi-sample problem is

$$T_{AB} = \max(1 - L_{T_A}, 1 - L_{T_B}),$$

where $T_A = \sum_{j=1}^C n_j (\bar{X}_j^*)^2$ and $T_B = \sum_{j=1}^C \left[\left(\sum_{i=1}^{n_j} I(X_{ij}^* > \tilde{M}) \right)^2 / n_j \right]$, $\bar{X}_j^* = \sum_{i=1}^{n_j} X_{ji}^* / n_j$. T_A

is a natural extension of T_a , because it is permutationally equivalent to the Snedecor's F -ratio statistic, while T_a is equivalent to the two-sample Student's t statistic (Pesarin 2001, Marozzi 2003). T_b is the same as T_a except computed on the dichotomous data transformation based on the pooled median, and so is T_B , which is equivalent to the F -ratio statistic computed on $I(X_{li}^* > \tilde{M})$ rather than on X_{li}^* . Marozzi (2003) showed that T_{ab} and T_{AB} have a similar good power behavior.

3. Application of Bi-Aspect Nonparametric Testing

In this section, three actual applications of bi-aspect testing are provided. Permutation tests were carried out from a random sample of 10000 permutations of test statistics.

We first apply the bi-aspect test to data from Foster et al. (1998). Two methods have been used to train service agents of a cellular telephone broadcast equipment company. The time in hours required to handle each of a sample of 150 calls for service was measured. These calls were randomly assigned to two groups, with 75 calls allotted to agents trained by the first method and the others to agents trained by the second method. The most common way to tackle this problem is the Student's t test that makes a crucial assumption about the underlying distribution, that of normality. As one might expect, since the data are waiting times, the assumption of normality is very difficult to justify. The histogram of the data shows that they are heavily skewed, and that some calls last considerably longer than others. Note that the presence of skewness and outliers affects the performance of the t test but not the performance of the bi-aspect test. The company would like to investigate about the performance of the methods in order to discard the one that leads to longer service times. We assume that the first method leads to longer service times than those of agents trained by the second method. The t test indicates that the two methods lead to comparable times, and so does T_a . In fact, their p -values are 0.2583 and 0.2553, respectively. On the contrary, T_{ab} (p -value 0.0101) indicates that the second method is better than the first one. This is due to the component T_b that produces a partial p -value of 0.0055.

The effects of cigarette smoking on maternal airway function during pregnancy were investigated by Das et al. (1991). The subjects were pregnant women. Measurements on forced vital capacity (FVC) were classified by smoking status of the women (non-, light- and heavy-smoker). We are interested in evaluating if FVC differs according to smoking status of the women. The Fisher's F and T_{AB} tests were used to answer this question. On the basis of our results, the null hypothesis that the population means are all equal is not rejected by the tests. However, it should be noted that the p -value of the Fisher and T_{AB} tests are rather different: the first one is about 60% and the other one is about 30%. T_A produces a p -value that is practically equal to that of the Fisher's test, whereas the p -value of T_B is about one third of these ones. This difference affects the p -value of their combination T_{AB} which is 0.2961.

A very common problem in quality control and industrial statistics is to determine whether two (or more) sample means are significantly different. Production lines in a

large industrial corporation are set to produce a specific type of steel ball bearing. Romano (1977) reported the measure of the diameter of ten ball bearings that were randomly picked from one production line and of ten ball bearings from another one. We wish to evaluate if there is practical evidence to reject the hypothesis of equal machine settings. For this reason, we test the null hypothesis that the mean diameter of the bearings of the two production lines is the same against the alternative hypothesis that bearings from the first line have a greater diameter than those from the second line. By inspecting the resulting p -values, we conclude that the data suggest there are no significant differences among the two settings. It should be noted that in this example application the p -values of T_a (10.15%) and T_b (32.43%) lead to a p -value of T_{ab} close to 12.7% that is very similar to that of the Student test which is close to 10.6%.

4. Conclusion

Marozzi (2002 & 2003) proposed two nonparametric tests for location problems, which are shown by simulation to have better power behavior than that of traditional parametric and permutation tests for location problems in the presence of outliers and highly skewed data. In this paper, a detailed and unified presentation of the idea and the theory of bi-aspect testing is given. Three actual applications in business, medical and industrial statistics are then discussed. In these examples, the bi-aspect testing allows us to draw deeper and more informative inference than that allowed by traditional methods.

References

- Das T. K., Moutquin J., Parent J. (1991) Effect of cigarette smoking on maternal airway function during pregnancy, *American Journal of Obstetrics and Gynecology*, 165, 675-679.
- Foster D. P., Stine R. A., Waterman R. P. (1998) *Basic business statistics*, Springer-Verlag, New York.
- Huber P. (1981) *Robust statistics*, Wiley, New York.
- Johnson N. L., Kotz S. (1970) *Continuous univariate distributions-I*, Houghton Mifflin, Boston.
- Marozzi M. (2002) A Bi-Aspect Nonparametric Test for the Two-Sample Location Problem, *Computational Statistics & Data Analysis*, forthcoming.
- Marozzi M. (2003) A Bi-Aspect Nonparametric Test for the Multi-Sample Location Problem, *Computational Statistics & Data Analysis*, forthcoming.
- Pesarin F. (2001) *Multivariate Permutation Tests with Applications in Biostatistics*, Wiley, Chichester.