0363-5465/97/2525-0585$02.00/0
THE AMERICAN JOURNAL OF SPORTS MEDICINE, Vol. 25, No. 4

Current Concepts

A Statistics Primer

Statistical Tests for Discrete Data

John E. Kuhn, MS, MD, Mary Lou V. H. Greenfield,* MPH, MS, and Edward M. Wojtys, MD

From MedSport and the Section of Orthopaedic Surgery, University of Michigan, Ann Arbor, Michigan

When reviewing articles in the literature, it is important to spend some time thinking specifically about the data presented. The data collected in an experiment may be classified into two general types. *Continuous data* represent data that vary over a continuous range and may be displayed on a curve, frequently a bell-shaped curve. Some examples of continuous data include height, age, and the times recorded in a mile race. *Discrete data,* on the other hand, are data that fall into specific categories. Examples of this type of data include gender, the presence or absence of a specific condition, and the classification of a patient's results into groups such as poor, fair, good, and excellent. Statistics provides different tests for analyzing these different types of data. In this review some of the commonly used tests for analyzing discrete data will be presented.

The frequency with which an observation occurs may be evaluated with the *chi-square test.* Suppose you read an article that reviews the frequency of three types of acromion shapes (flat, curved, or hooked) in 33 cadaveric specimens with rotator cuff tears. The null hypothesis would state that the acromion shape has no relation to rotator cuff tears and the expected or theoretical frequency of specimens with rotator cuff tears should be the same with each type of acromion; that is, 11 specimens should be found in each of the 3 groups. The data that were collected, however, show a higher incidence of rotator cuff tears in specimens with curved acromiae, and an even higher incidence in specimens with hooked acromion types, with 3 specimens with flat acromiae, 8 with curved acromiae, and 22 with hooked acromiae.

Is this trend significant? You would like to know a P value to assess the probability that these observations represent the truth and are not due to chance.

The chi-square test is used to evaluate these discrete data and to derive a P value. The chi-square test uses an equation that includes the observed frequency of an occurrence, the expected or theoretical frequency of an occurrence, and the number of groups; the equation is used to generate a number called the chi-square statistic (Fig. 1). The investigator may find this number in a standard table in the back of any statistics book and determine the P value for the significant difference between the observed and theoretical frequencies in the groups. This P value depends on the degrees of freedom in the experiment, which, in general, is the number of groups of observations minus one. In this example, with three types of acromion shapes, there are two degrees of freedom.

There are many ways to misuse this valuable test.[2] One of the more common misuses of the chi-square test is to use small sample sizes or small theoretical frequencies. For the chi-square test to be applied correctly, the generally recommended sample size of the total number of occurrences in all groups should be greater than 50, and each group should have at the very minimum 5—and preferably 10—occurrences. These recommendations are only guidelines; however, in circumstances where these recommendations are not met, there are modifications or corrections of the chi-square test that allow for meaningful and more accurate chi-square analysis. Unless these modifications or corrections are used, investigators may create errors in the level of significance when evaluating their data.

For example, when two groups are being compared (the degree of freedom is one) and the expected or theoretical frequency for any of the groups is less than 10, the chi-square test may produce an overestimation of the chi-square statistic, leading to an error and an overestimation of the significance of the difference between the groups. (This occurs as a result of approximations made in the

* Address correspondence and reprint requests to Mary Lou V. H. Greenfield, MPH, MS, University of Michigan, Orthopaedic Surgery, TC2914G-0328, 1500 East Medical Center Drive, Ann Arbor, MI 48109.

No author or related institution has received any financial benefit from research in this study.

$$\chi^2 = \sum_{i=1}^{a} \frac{(f_i - f_j)^2}{f_j}$$

where

f_i = the observed frequency of the ith group

f_j = the expected frequency of the jth group

a = the number of groups

Example

Frequency	Acromion Morphology		
	Flat	Curved	Hooked
Observed	3	8	22
Expected	11	11	11

Then, using the χ^2 equation,

$$\chi^2 = \frac{(3-11)^2}{11} + \frac{(8-11)^2}{11} + \frac{(22-11)^2}{11}$$

$$\chi^2 = 5.818 + 0.818 + 11$$

$$\chi^2 = 17.636$$

In the table, with 2 *df*, the *P* value for this χ^2 statistic is less than 0.005.

Figure 1. Example of a chi-square test: the Pearson chi-square example is one of the most commonly used chi-square tests.

derivation of the original chi-square formula.) In this situation, the *Yates' correction for continuity* may be applied, which generates a more conservative chi-square statistic and makes the *P* value more accurate.[3]

The *Fisher's exact test* is another useful modification of the chi-square test. This test is more conservative as it corrects for smaller numbers of observations and is used when the numbers of observations in a given grouping are less than 10 and the data may be arranged into a 2 × 2 table.

Another common error made by some investigators is the inappropriate use of the chi-square test when there is not independence among the groups. Two or more events are said to be independent of one another if the probability of an event occurring is the same regardless of whether the other event has occurred or if that event has not occurred.[1] Consider, again, the example of acromion shape and rotator cuff tears. Suppose that the natural development of an acromion is for it to begin as a flat structure at birth, gradually progress to a curved shape in middle life, then change to a hooked shape in later life. The probability, then, that a patient has one of three types of acromion shape (flat, curved, or hooked) *is not independent* of the distinct age group (young age, middle age, older age) in which they fall. This means, for example, that a greater proportion of older people will have hooked acromiae than younger people. If we studied acromion shape in our cadaveric model and did not take into consideration the effect of age group of the specimen on morphology type, we would get the wrong answer when using a chi-square test.

Finally, the data analyzed by the chi-square test must be discrete; if continuous data are used the chi-square statistic may be distorted in one of two ways. First, if one were to measure height (a continuous variable) in boys and girls and apply a chi-square test to detect significance, the resulting table would be huge; instead of a 2 × 2 table, or a 2 × 10 table, the table might be a 2 × 50 table because height is a continuous variable and each different measurement would result in a different row in the table. According to sample size requirements discussed previously, a huge sample size would be necessary to fill all of the cells of a table. Second, the chi-square statistic varies with the size of the units employed in the measurement. Consider measuring height in either inches or centimeters among boys and girls; the chi-square statistic for measuring height in centimeters would be 2.54 times the same chi-square statistic measuring height in inches. Clearly, the results would lead to erroneous conclusions regarding the differences between the height of girls and boys in this example. Statistical tests for continuous data are the appropriate choice for such measurements and will be presented in a future review.

ACKNOWLEDGMENT

The authors thank Dr. M. Anthony Schork from the University of Michigan, School of Public Health, Department of Biostatistics, for his review of this article.

REFERENCES

1. Hirsch RP, Riegelman RK: *Statistical First Aid: Interpretation of Health Research Data.* Boston, Blackwell Scientific Publications, 1992, p 9
2. Lewis D, Burke CJ: The use and misuse of the chi-square test, in Lieberman B (ed): *Contemporary Problems in Statistics. A Book of Readings for the Behavioral Sciences.* New York, Oxford University Press, 1971, pp 280–319
3. Yates F: Contingency tables involving small numbers and the χ^2 test. *J Royal Stat Soc (Suppl) 1:* 217–235, 1934