International Journal of Research in Science and Technology
(IJRST) 2013, Vol. No. 3, Issue No. I, Jan-Mar

http://www.ijrst.com
ISSN: 2249-0604

STATISTICAL HYPOTHESIS TESTING & LINEAR CORRELATION

Devkatte Patil Sandip Marotrao

Dept. of Mathematics, Cmj University, Shillong, Meghalaya

INTRODUCTION

The hypothesis originally proposed by Galton and elaborated by Spearman, that there is a functional correspondence between sensory discrimination and general intelligence (*g*) continues to spark debate. Previous findings suggest that pitch discrimination and tactile discrimination are only weakly correlated with *g*. This study sought to replicate the pitch discrimination findings and to expand them to the modality of color discrimination in a large sample (N = 899) by correlating two sensory discrimination measures with the general factor from a battery of 13 cognitive-ability tests. The modest correlations found between *g* and measures of pitch discrimination (r = .21) and color discrimination (r = .31) suggest that sensory discrimination is relatively distinct from general intelligence. Although consistent with the neural processing speed explanation of *g*, these results cast doubt on a strong form of the sensory discrimination explanation of *g*.

Sensory discriminative analysis forms a fundamental type of methodology and is used widely in sensory and consumer research. Sensory Discrimination Tests and Measurements: Statistical Principles, Procedures and Tables provides a comprehensive discussion of sensory discriminative analysis from a statistical perspective. Designed to be both a reference manual and a research monograph, practitioners will discover various useful test and measurement procedures. More statistically oriented readers will find the statistical considerations behind the procedures. Sensory Discrimination Tests and Measurements will be of interest to everyone concerned with testing and measuring sensory difference and consumer preference.

Discrimination testing is a technique employed in sensory analysis to determine whether there is a detectable difference among two or more products. The test uses a trained panel to discriminate from one product to another.

Statistical basis

The statistical principle behind any discrimination test should be to reject a null hypothesis (H_0) that states there is no detectable difference between two (or more) products. If there is sufficient evidence to reject H_0 in favour of the alternative hypothesis, H_A: There is a detectable difference, then a difference can be recorded. However, failure to reject Ho should not be assumed to be sufficient evidence to accept it. H_0 is formulated on the premise that all of the assessors guessed when they made their response. The statistical test chosen should give a probability value that the result was arrived at through pure guesswork. If this probability is sufficiently low (usually below 0.05 or 5%) then H_0 can be rejected in favour of H_A.

http://www.ijrst.com
ISSN: 2249-0604

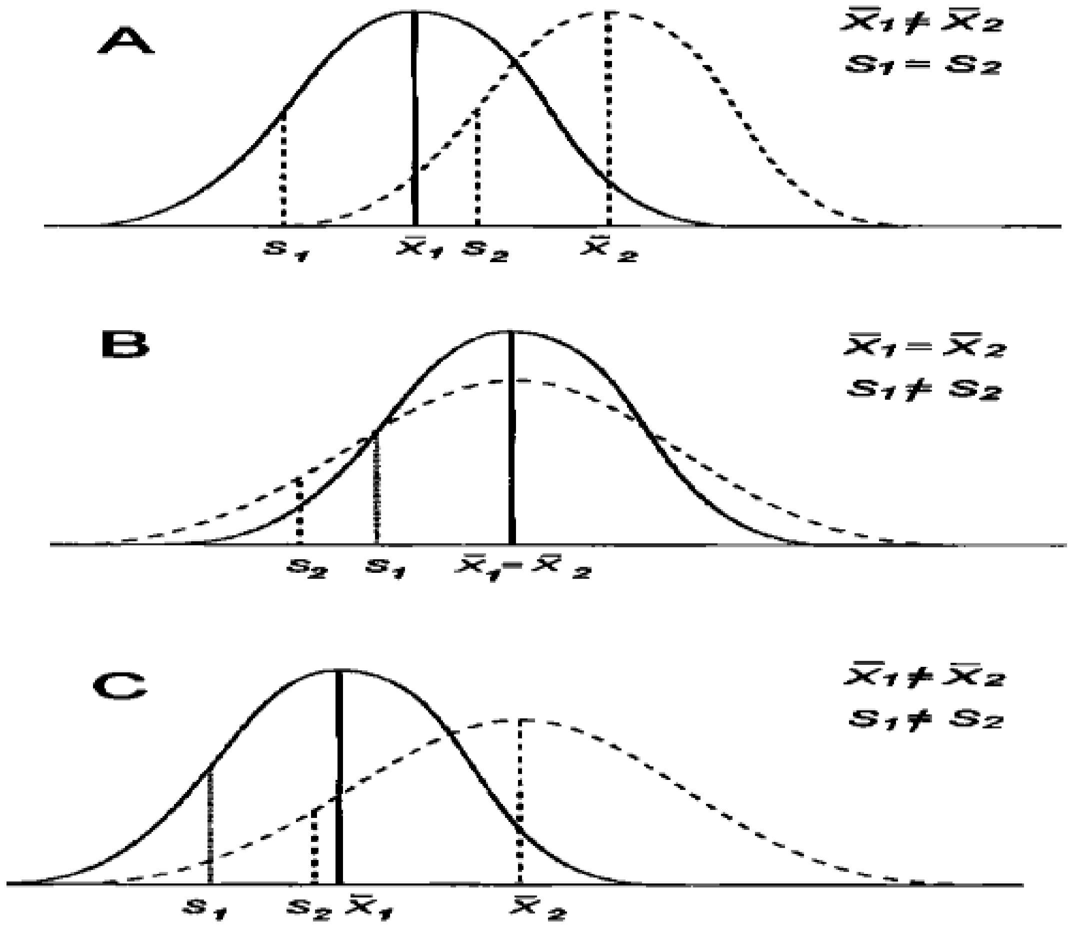
REVIEW OF LITERATURE

In analytical work a frequently recurring operation is the verification of performance by comparison of data. Some examples of comparisons in practice are:

- Performance of two instruments,
- Performance of two methods,
- Performance of a procedure in different periods,
- Performance of two analysts or laboratories,
- Results obtained for a reference or control sample with the "true", "target" or "assigned" value of this sample.

Some of the most common and convenient statistical tools to quantify such comparisons are the F-test, the t-tests, and regression analysis.
Because the F-test and the t-tests are the most basic tests they will be discussed first. These tests examine if two sets of normally distributed data are similar or dissimilar (belong or not belong to the same "population") by comparing their standard deviations and means respectively. This is illustrated in Fig. 1

Fig. 1. Three possible cases when comparing two sets of data ($n_1 = n_2$). A. Different mean (bias), same precision; B. Same mean (no bias), different precision; C. Both mean and precision are different. (The fourth case, identical sets, has not been drawn).



http://www.ijrst.com
ISSN: 2249-0604

Two-sided vs. one-sided test

These tests for comparison, for instance between methods A and B, assume that there is no significant difference (the "null hypothesis"). In other words, when the difference is so small that a tabulated critical value of F or t is not exceeded, we can be confident (usually at 95% level) that A and B are not different. Two fundamentally different questions can be asked concerning both the comparison of the standard deviations s_1 and s_2 with the F-test, and of the means $\bar{x}_1$, and $\bar{x}_2$, with the t-test:

1. Are A and B different? (two-sided test)
2. Is A higher (or lower) than B? (one-sided test).

This distinction has an important practical implication as statistically the probabilities for the two situations are different: the chance that A and B are only different ("it can go two ways") is twice as large as the chance that A is higher (or lower) than B ("it can go only one way"). The most common case is the two-sided (also called two-tailed) test: there are no particular reasons to expect that the means or the standard deviations of two data sets are different. An example is the routine comparison of a control chart with the previous one (see 8.3). However, when it is expected or suspected that the mean and/or the standard deviation will go only one way, e.g., after a change in an analytical procedure, the one-sided (or one-tailed) test is appropriate. In this case the probability that it goes the other way than expected is assumed to be zero and, therefore, the probability that it goes the expected way is doubled. Or, more correctly, the uncertainty in the two-way test of 5% (or the probability of 5% that the critical value is exceeded) is divided over the two tails of the Gaussian curve, i.e., 2.5% at the end of each tail beyond 2s. If we perform the one-sided test with 5% uncertainty, we actually increase this 2.5% to 5% at the end of one tail. (Note that for the whole gaussian curve, which is symmetrical, this is then equivalent to an uncertainty of 10% in two ways!)

This difference in probability in the tests is expressed in the use of two tables of critical values for both F and t. In fact, the one-sided table at 95% confidence level is equivalent to the two-sided table at 90% confidence level.

It is emphasized that the one-sided test is only appropriate when a difference in one direction is expected or aimed at. of course, it is tempting to perform this test after the results show a clear (unexpected) effect. In fact, however, then a two times higher probability level was used in retrospect. This is underscored by the observation that in this way even contradictory conclusions may arise: if in an experiment calculated values of F and t are found within the range between the two-sided and one-sided values of F_{tab}, and t_{tab}, the two-sided test indicates no significant difference, whereas the one-sided test says that the result of A is significantly higher (or lower) than that of B. What actually happens is that in the first case the 2.5% boundary in the tail was just not exceeded, and then, subsequently, this 2.5% boundary is relaxed to 5% which is then obviously more easily exceeded. This illustrates that statistical tests differ in strictness and that for proper interpretation of results in reports, the statistical techniques used, including the confidence limits or probability, should always be specified.

http://www.ijrst.com
ISSN: 2249-0604

METHOD OF RESEARCH

At least four reasons account for Spearman's lasting and even growing renown:

(1) His theoretical and methodological contributions were notably creative and original and have been continually used and developed further over the last fifty years.

(2) They have engendered controversy, theoretical arguments, and an increasing rate of empirical research testing his theories.

(3) The chief topic of his research – the nature and causes of variation in human intelligence– has attracted enormous interest with the spread of universal public education, the visibly increased range of individual differences in educability and scholastic performance in the school-age population, and the increasing cognitive requirements for employability in the technological, information-intensive modern world.

(4) Finally, Spearman established a coherent scientific approach to the study of variation in human behavioral traits that came to be known as the London School, so named because his work and influence were associated with his position as professor (and head)) of the Department of Psychology in University College, London. The London School's essentialaim is the advancement of psychology as an empirical, quantitative, biological science.Specifically, it has focused on the measurement and taxonomy of individual differences in human mental abilities and personality traits and the investigation of their nature and nurture, that is to say, the genetic and environmental factors responsible for the wide range ofdifferences between individuals in all psychological or behavioral characteristics. Researchers pursuing the aims of the London School are inescapably indebted to Spearman's pioneer efforts. Moreover, even today they continue to find in his works many key insights and hypotheses for empirical investigation.

CONCLUSION

For the initial error scores on the 100-Hue (prior to the square-root transformations), the mean score was 60.47, with a standard deviation of 50.79. This distribution is comparable to the general-population distribution reported in the test manual (Farnsworth, 1957). After transformations, the mean score was 44.88, with a standard deviation of 28.56. Correlations between the 100-Hue and the standard-battery tests, excluding Pitch Discrimination ranged from .08 to .32 (M = .20, SD = .07). The 100-Hue's highest correlations were with Paper Folding and Color Perception. The correlation with Color Perception shows that the two tests measure related yet distinct traits.

The mean score on Pitch Discrimination was 62.22, with a standard deviation of 9.62. This distribution is comparable to the general-population distribution reported by Seashore et al.

(1940). Correlations between Pitch Discrimination and the other standard-battery tests ranged from .03 to .49 (M = .16, SD = .11). Pitch Discrimination's highest correlations were with the two other auditory tests, Tonal Memory and Rhythm Memory. These correlations show that these tests also measure related yet distinct traits.

The analyses using a hierarchical model of cognitive abilities yielded four first-order factors, with initial eigenvalues (before rotation) of 4.52, 1.28, 1.11, and 1.04, and one second-order factor (g) In the oblique rotation, we allowed the first-order factors to have modest intercorrelations, ranging from .10 to .33. The first first-order factor showed its highest loadings on spatial tests such as Paper Folding, Memory for Design, and Wiggly Block and was interpreted as an index of structural visualization. The second first-order factor showed its highest loadings on memory tests such as Number Memory and Silograms (verbal-associative memory) and was interpreted as an index of memory. The third first-order factor showed its highest loadings on tests such as Number Series, Number Facility, English Vocabulary, and Analytical Reasoning, which are related in content to skills learned in school and was interpreted as an index of academic ability. The fourth first-order factor showed its highest loadings on tests such as Inductive Reasoning and Observation, which involve quickly noticing visual features, and was interpreted as an index of rapid visual processing. The fifth first-order factor, with an initial eigenvalue of only .88, was not retained in the rotated solution.

The first second-order factor showed high loadings on each of the lower-order factors (.69, .70, .57, and .53, respectively) and was interpreted as g. It showed substantial correlations with nearly all of the cognitive-ability tests (M = .57, SD = .13). The second second-order factor, with an eigenvalue of only .72, was disregarded.

The 100-Hue was only moderately related to general intelligence. When the effects of age and gender were controlled, the correlation between the 100-Hue and g was .31. Pitch Discrimination showed a somewhat lower partial correlation of .21 with g.

REFERENCES

Bethscheider, J. K. (1990). The color discrimination project (Technical Report 1990-2). Chicago: Johnson O'Connor Research Foundation.

Cattell, J. M. (1886a). The inertia of the eye and brain. Brain, 8, 295-381.

Daniel, M. H. (1982). A factorial study of reasoning tests (Technical Report 1982-6). Boston: Johnson O'Connor Research Foundation. (ERIC Document Reproduction Service No. ED 230 571)

Farnsworth, D. (1957). The Farnsworth-Munsell 100-Hue Test for the examination of color discrimination: Manual (rev. ed.). Baltimore: Macbeth.

Li, S-C., Jordanova, M., & Lindenberger, U. (1998). From good senses to good sense: A link between tactile information processing and intelligence. Intelligence, 26, 99-122.

Raz, N., Willerman, L., & Yama, M. (1987). On sense and senses: Intelligence and auditory information processing. Personality and Individual Differences, 8, 201-210.

Spearman, C. (1904). "General intelligence," objectively determined and measured. American Journal of Psychology, 15, 201-293.

Tal, J. S. (1986). Aptitudes of guidance counselors (Technical Report 1986-2). Chicago: Johnson O'Connor Research Foundation. (ERIC Document Reproduction Service No. ED 289 903)

Wissler, C. (1901). The correlation of mental and physical tests. Psychological Review, Monograph No. 3.

Zimowski, M. F., & Wothke, W. (1988). The measurement of structural visualization: An evaluation of spatial and nonspatial sources of variation in the Wiggly Block and Paper Folding test scores (Technical Report 1988-5). Chicago: Johnson O'Connor Research Foundation. (ERIC Document Reproduction Service No. ED 305 384)