Directional and nondirectional alternative hypotheses

JOHN GAITO

York University, Downsview M3J 1P3, Ontario, Canada

The problem of one- vs two-tailed tests is discussed. The use of the terms "one tailed" and "two tailed" can lead to confusion, for example, with ANOVA procedures. It is suggested that these terms be replaced with "directional" and "nondirectional" alternative hypotheses. Furthermore, directional alternative hypotheses appear to contradict the operation of the null hypothesis. Thus, it appears that nondirectional alternative hypotheses should be used in all cases.

Some years ago the controversy concerning the use of one-tailed tests of statistical hypotheses was quite prominent, for example, Burke (1953), Goldfried (1959), Jones (1952, 1954), and Kimmel (1957). This controversy has largely subsided, although the problem still has not been resolved; there persists in the literature frequent use of one-tailed tests. The purpose of this paper is to discuss two suggestions relative to this issue: (1) The use of the terms "one tailed" and "two tailed" can lead to confusion; the terms "directional" and "nondirectional" alternative hypotheses would be more appropriate. (2) Directional statistical hypotheses contradict the operation of the null hypothesis. Thus, nondirectional alternative statistical hypotheses should be used in all cases.

BACKGROUND

To illustrate this problem, let us take as an example an experiment involving two groups, Group A and Group B. The experimental hypothesis is that $A > B$ ($\mu_a > \mu_b$, using parameters). To evaluate this empirical hypothesis, the researcher may use the t distribution. This distribution is one which results when two random samples are chosen independently from a single normal distribution and the difference between the two means is determined. The mean difference is zero under the null hypothesis. This distribution involves only random differences (random errors), with positive random errors as likely as negative errors. This is often stated as the es are NID (0, σ_e^2), that is, the errors are normally independently distributed, with a mean of zero and a variance, σ_e^2. The null hypothesis (H_0) is stated as $\mu_a = \mu_b$ or $\mu_a - \mu_b = 0$. In a two-tailed case (nondirectional), the alternative hypothesis would be $\mu_a \neq \mu_b$. In a one-tailed situation (directional), H_0 would be the same as the above but the alternative hypothesis would be $\mu_a > \mu_b$ or $\mu_a - \mu_b > 0$ for the above empirical hypothesis ($A > B$). If the empirical hypothesis was $B > A$, the alternative hypothesis would be $\mu_b > \mu_a$ or $\mu_b - \mu_a > 0$.

"ONE-TAILED" AND "TWO-TAILED" TERMS CAN BE MISLEADING: SUBSTITUTE THE TERMS "DIRECTIONAL" AND "NONDIRECTIONAL" ALTERNATIVE HYPOTHESES

A discussion of the problem when using a t test is quite simple. The usual one-tailed alternative is to be found in the right tail or left tail of the t distribution. However, when we use ANOVA procedures with resulting F tests, confusion can occur. This aspect is illustrated in some statistical textbooks (e.g., Horowitz, 1974, p. 285) which state that F tests are "one tailed." This is a factual statement relative to the H_0 rejection region within this distribution. However, this is an irrelevant point relative to the one-tailed vs two-tailed issue. In terms of alternative hypotheses, the ANOVA procedures are nondirectional ("two tailed"). An estimate of μ is obtained (general mean) and the estimate means of each treatment (μ_i) go in both directions, that is, some estimated means are greater than the estimate of μ and others are less than this estimate. Thus, one is faced with an anomalous situation in which he would assert that ANOVA procedures are both "two tailed" and "one tailed" in the same situation. They are "two tailed" inasmuch as a nondirectional H_1 is involved; however, the resulting F test is "one tailed" in terms of the rejection region of concern. This confusion can be avoided if we discard the terms "one-tailed" and "two-tailed" tests and substitute "directional" and "nondirectional" alternative hypotheses.

DIRECTIONAL ALTERNATIVE HYPOTHESES CONTRADICT THE OPERATION OF THE NULL HYPOTHESIS: NONDIRECTIONAL HYPOTHESES SHOULD BE USED

In both directional and nondirectional cases, the null hypothesis is the same, that is, $\mu_a = \mu_b$ or $\mu_a - \mu_b = 0$; it is only the H_1s which are different. Furthermore, in both cases we are concerned with *tests of the null hypothesis*. Also, in these statistical tests of H_0, the same

distribution or comparator curve is used, for example, the t distribution, a distribution which occurs when H_0 is true. Thus, the emphasis seems to be on the null hypothesis.

When a specific curve such as the t distribution is chosen, the purpose is to contrast the observed results with those expected when H_0 is true. This distribution indicates that random negative errors (in the left tail) are as frequent as random positive errors (in the right tail). The t distribution is not a directional one under H_0.

If the investigator decides to make use of only one of the two tails of the t distribution, it appears as if he is using only one-half of the distribution. This seems to run counter to the concept under which this distribution operates, that is, that positive differences are as frequent as negative between the means. Such procedure seems to contradict the operation of the null hypothesis.

This whole problem is resolved automatically with ANOVA procedures. In this case there are no directional alternatives. The alternative hypothesis is $\mu_a \neq \mu_b$, if only two groups are involved, or $\Sigma(\mu_i - \mu) = 0$, if more than two groups occur (μ_i is the mean of the ith group and μ is estimated by all of the data). There is a single estimate of μ, and there are μ_is which are greater than this estimate and some which are less than this estimate, such differences reflecting random error (under H_0). Thus, although the central F distribution is "directional" in terms of the rejection region, the difference between means, which is the basis for the sums of squares and ultimately the mean squares for the numerator of the F test, allows for positive and negative differences around the estimate of μ. This nondirectional aspect is reflected in the familiar relationship between t and F, that is, $t^2 = F$, which holds only when the t is nondirectional (i.e., "two tailed").

The fact that ANOVA procedures use nondirectional alternative hypotheses seems to reinforce the thesis that directional alternatives are to be avoided.

ROLE OF DIRECTIONAL HYPOTHESES

Given the above theses, one may wonder how directional alternative hypotheses enter the picture. It seems that these hypotheses are intimately intertwined with the empirical hypotheses of the researcher and should be of concern in the interpretation of the results.

For example, let us assume that with α set at a probability value of .05, H_0 has been rejected and that the mean for Group A is significantly greater than the mean for Group B. Based on the fact that H_1 has two possibilities ($\mu_a > \mu_b$ and $\mu_b > \mu_a$), each with a p of .025, the investigator might accept $\mu_a > \mu_b$ as the true state of events, if it is consistent with his empirical hypotheses or other logical considerations. This is an inferential process based on results of tests of statistical hypotheses but relying on interpretation by the investigator and is directional in nature.

REFERENCES

BURKE, C. J. A brief note on one-tailed tests. *Psychological Bulletin*, 1953, **50**, 384-387.

GOLDFRIED, M. R. One-tailed tests and "unexpected" results. *Psychological Review*, 1959, **66**, 79-80.

HOROWITZ, L. M. *Elements of statistics for psychology and education*. New York: McGraw-Hill, 1974.

KIMMEL, H. D. Three criteria for the use of one-tailed tests. *Psychological Bulletin*, 1957, **54**, 351-353.

JONES, L. V. Tests of hypotheses: One-sided vs two-sided alternatives. *Psychological Bulletin*, 1952, **49**, 43-46.

JONES, L. V. A rejoinder on one-tailed tests. *Psychological Bulletin*, 1954, **51**, 585-586.