Statistical significance
Martin Kozloff

Statistical significance means that an outcome was not likely the result of chance--i.e., a random distribution. Outcome might mean the average (mean) scores at the post test of an experimental vs. comparison group; or the mean pre vs post test score of one group; of the percentage of students passing end of grade tests in two or more schools.

What makes an outcome statistically significant?

If an immortal monkey were typing on a typewriter with thirty keys (letters, period, space, comma, semi-colon), what are the odds that he (it's a male monkey) would ONE time type Hamlet perfectly? Well, what is the first keystroke? W. What are the odds of W? 1/30. What is the next keystroke? O. What are the odds of O. 1/30. What are the odds of WO. 1/30x30. So, if Hamlet has a million keystrokes, what are the odds of the monkey typing Hamlet one time? 1/30 to the millionth power.

Does this ONE perfect Hamlet mean that it was NOT random--that it was instead an intelligent act; the monkey knew what he (it's a male monkey, okay!!) was doing After all, the odds are so small---1/30 to the millionth power.

NOOOO. Typing Hamlet does NOT beat the odds because the odds of ANY sequence of 1,000,000 keystrokes is 1/30 to the millionth power.

But what if Zippy does Hamlet 2 times in a row???!!! What are the odds of THAT????!!!! Now, THAT would be "against the odds"--statistically significant.
1/30 to the millionth power x 1/30 to the millionth power.

A more ralistic example

Okay, so now you have two groups: traditional math instruction vs. Personalized System of Instruction. Your research hypo is that there is higher achievement with PSI.

Let's say that there are 20 students in each group. And let's say that the average score on the pre-test was 1--out of a possible 6. And let's say that the average post test score was 5 for the PSI and 2 for traditional.

**What is the NULL hypothesis?.... The difference in scores (outrcome) is not stat signif--easily chance.

Well, then, **the test (of the null hypo) is the stat test of whether these differences are or are not ss. It's all a matter of probability.

Imagine that you have one die in each hand--left = traditional and right = PSI. You roll die for trad student one and get a 2. You roll the die for PSI student one and get a five. What are the odds of that? Well, how many combinations are there? 6 x 6 = 36. So, the odds of 2 and 5 are 1/36.

You roll the dice for student two in each group. Trad = 1; PSI = 6. What are the odds of THAT? The odds that the scores for BOTH students are in a direction that supports the research hypothesis?

You roll again. 2 and 5
Again, 1 and 6.

In the end, you have 18 students in the trad group with post test scores of 3 or less; you have 19 students in the PSI group with post test scores 4 or more. What are the odds of THAT? If the dice are NOT loaded (i.e., if the results WERE chance) the distribution of scores (so many 1s, 2s, 3s, etc., for each group) would be about equal. But they are WAY diferent from equal. The scores are heavily skewed towards high scores for PSI and low scores for trad.

There are tables that tell you, if you have x number of participants, and each one has a certain scores, here are the odds that you would get mean differences between two groups 1/10,00 0 by chance, 1/1000 times, 1/100 times, etc.

In general, when the odds are less than 5/100, we say, "That is statistically significant.

Note, even small differences (between groups or within one group pre to post) are stat signif.

But a large diff in a small group may NOT be ss. Why? Becaause one or two flukey scores will make a big diff in the MEAN.

Correlation

Correlation has to do with how much knowledge of the value of one instance of a variable (e.g., one child's age) enables you to predict the value of another variable about that child (e.g., weight) better than if you knew the average score of a group. If the average weight of a sample were 90 pounds and, not knowing anything else about Billy you guessed that his weight was 90, you might be wrong. his weight (120) - mean (90) = 30. This is error. But what if you know that food intake was associated with weight positively. Now, if you know how many calories fat boy takes in, you might predict that his weight is 115. Now your error is only 5.
ficient

So, a **correlation coefficient** is a number gotten by going through an algorithm that tells you how well one variable predicts another