HOW TO ....
A PRACTICAL GUIDE TO PSYCHOMETRICS

Calculate the probability that two scores are reliably different

Contents

If viewing on screen you can click on a contents item above to jump to the page the item is on

1. When you have two scores for one individual on different tests or scales

Use the Z-Score formula for this. To find out whether there is a reliable difference between two scores (which might of course be subtest scores) calculate the value of:

$$Z_{difference} = \frac{Z_x - Z_y}{\sqrt{2 - (r_{xx} + r_{yy})}}$$

The probability of the difference being due to errors of measurement can be found by looking up the probability corresponding to the value of $Z_{difference}$ in tables for the normal curve.

Example. Someone obtains a Verbal IQ of 118 and a Performance IQ of 106 on an intelligence test with a mean of 100 and a standard deviation of 15 for all IQ,s. What is the probability that this difference is due to chance? The reliability of the Verbal Scale is .84 and of the Performance Scale is .80. $Z_{difference}$ will therefore equal:

$$Z = \frac{(1.2 - 0.4)}{\sqrt{2 - (0.84 + 0.80)}} = \frac{0.8}{\sqrt{2 - 1.64}} = \frac{0.8}{0.6} = 1.25$$

A $Z_{difference}$ of 1.25 yields a 2-tail p value of .21. If we adopt the conventional probability levels for deciding that a result is ‘significant; then this one is not. We would have to conclude that there is insufficient reason for believing that the true score for Verbal IQ is any higher than the true score for Performance IQ.

2. When you have two scores for one individual on the same test or scale

This situation arises when somebody is retested.

The formula is very similar to the one above. The difference is that in the above formula we find the reliabilities of each of the two tests. In the present case we only have one test. The formula is:

$$Z_{difference} = \frac{Z_{x_1} - Z_{x_2}}{\sqrt{2 - 2r_{xx}}}$$

If the test is one on which a practice effect can be expected the formula should be modified as follows. The test given second will need to have the practice effect subtracted from its value, thus.

$$Z_{difference} = \frac{Z_{x_1} - Z_{(x_2 - p)}}{\sqrt{2 - 2r_{xx}}}$$

where *p* is the practice effect

Example. A person scores at the 95th percentile on a test of depression. On retest following treatment the score is at the 84th percentile. If the reliability coefficient of the test is .82, what is the probability that there has been a change in true Depression score?

Answer. Applying the formula we get the following values

$$Z_{difference} = \frac{1.64 - 1.00}{\sqrt{2 - (2 \times 0.82)}} = \frac{0.64}{0.60} = 1.07$$

Consulting Tables for the normal distribution we find that about 28 percent of people would show a larger difference than this on a chance basis, and that about 14 percent would show a drop in score as great as this. So there is no strong evidence to suppose that this person's true depression score has changed.

3. When you have two scores for two individuals on the same test or scale

This situation arises when you compare two individuals on their scores on the some scale and need to know the likelihood that their scores differ from each other, taking into account errors of measurement on the test.

The formula is the same as used above. The difference is that in this case we are comparing the scores of two individuals rather than the same individual tested on two occasions.

The formula is:

$$Z_{difference} = \frac{Z_{x_1} - Z_{x_2}}{\sqrt{2 - 2r_{xx}}}$$

Example: Two candidates for a job are given the same IQ test as part of the selection process. Candidate A obtains an IQ score of 124, candidate B obtains an IQ score of 127. The mean of the test is 100, the standard deviation 15 and the reliability is 0.91. Is the difference in their scores an indicator of real differences in their abilities or is it likely to that it is a result of error of measurement associated with the test?

Answer. Applying the formula we get the following values

$$Z_{difference} = \frac{1.80 - 1.60}{\sqrt{2 - (2 \times 0.91)}} = \frac{0.20}{0.42} = 0.48$$

Consulting Tables for the normal distribution we find that the probability of obtaining this difference or larger between two individuals on this test is about 63 percent on a chance basis.

Therefore the difference in test scores is not likely to be indicating differences in ability.