Statistical Reliability [1]

Statistical reliability is needed in order to ensure the validity and precision of the statistical analysis.

It refers to the ability to reproduce the results again and again as required. This is essential as it builds trust in the statistical analysis and the results obtained.

> For example, suppose you are studying the effect of a new drug on the blood pressure in mice. You would want to do a number of tests and if the results are found to be good in controlling blood pressure, you might want to try it out in humans too.
> The statistical reliability is said to be low if you measure a certain level of control at one point and a significantly different value when you perform the experiment at another time.

However, if the reliability is low, this means that the experiment [3] that you have performed is difficult to be reproduced with similar results then the validity of the experiment decreases. This means that people will not trust in the abilities of the drug based on the statistical results you have obtained.

Validity and Reliability

In many cases, you can improve the reliability by taking in more number of tests and subjects. Simply put, reliability is a measure of consistency.

Reliability can be measured and quantified using a number of methods.

Consider the previous example, where a drug is used that lowers the blood pressure in mice. Depending on various initial conditions, the following table is obtained for the percentage reduction in the blood pressure level in two tests. (Disclaimer: This is just an illustrative example - no test has actually been conducted)

Time after injection	Test 1	Test 2
1 min	5.86	5.89
2 min	6.35	6.41

3 min	7.12	6.95
4 min	9.18	9.01
5 min	12.36	12.13
6 min	14.26	14.93
7 min	16.96	15.89

Ideally, the two tests should yield the same values, in which case the statistical reliability will be 100%. However, this doesn't happen in practice, and the results are shown in the figure below. The dotted line indicates the ideal value where the values in Test 1 and Test 2 coincide.

Statistical Reliability

Image not found or type unknown

Using the above data, one can use the change in mean, study the types of errors in the experimentation including Type-I and Type-II errors or using retest correlation [4] to quantify the reliability.

The use of statistical reliability is extensive in psychological studies, and therefore there is a special way to quantify this in such cases, using Cronbach's Alpha [5]. This gives a measure of reliability [6] or consistency [7].

With an increase in correlation between the items, the value of Cronbach's Alpha increases, and therefore in psychological tests and psychometric studies, this is used to study relationship between parameters and rule out chance processes.

Source URL: https://staging.explorable.com/statistical-reliability

Links
[1] https://staging.explorable.com/statistical-reliability
[2] https://staging.explorable.com/
[3] https://staging.explorable.com/experimental-research
[4] https://staging.explorable.com/statistical-correlation
[5] https://staging.explorable.com/cronbachs-alpha
[6] https://staging.explorable.com/definition-of-reliability
[7] https://staging.explorable.com/internal-consistency-reliability