

How many replications?

<http://marekrychlik.com/node/14>

{The factors determining the number of replications:

- {The variance σ^2
- {The size of the difference that has physical significance, between two means δ .
- {The significance level of the test α , or the probability of Type I error.
- {The power of the test $1 - \beta$ (the probability of detecting difference δ) where β is the probability of Type II error.

{The required replication number for each treatment group r , for two-sided alternatives, can be estimated with the following formula:

$$r \geq 2 [z_{\alpha/2} + z_{\beta}]^2 \left(\frac{\sigma}{\delta}\right)^2 .$$

where $z_{\alpha/2}$ is the value of the standard normal variate exceeded with probability $\alpha/2$ and z_{β} is exceeded with probability β . Typical values are $\alpha = 0.05$ and $\beta = 0.2$, yielding the power of the test of 0.8. Let us comment about the Stevens measurement classification:

- {nominal (also categorical or discrete)
- {ordinal
- {interval
- {ratio

{The formula (??) applies to interval measurements, such as temperature in the Fahrenheit or Celsius scale. When we are dealing with a ratio measurement, such as temperature in the Kelvin scale, we may modify the above discussion by using relative analogues of σ and δ . We assume that the expected value of the measurement is μ on the ratio scale. The percent coefficient of variation is defined as follows:

$$\%CV = 100 \cdot \left(\frac{\sigma}{\mu}\right)$$

where σ is the true standard deviation and μ is the true expected value of the mean. Similarly, we may define the measure of the difference as percentage of the overall expected mean:

$$\%\delta = 100 \cdot \left(\frac{\delta}{\mu}\right)$$

The values $\%CV$ and $\%\delta$ may be used in place of σ and δ in formula (??) if one prefers the ratio scale. The factors that result in increase of r ;

- {The variance, $\%CV$ or σ^2 increases.
- {The size of the difference, $\%\delta$ or δ , decreases.
- {The significance level of the test α decreases.
- {The power of the test $1 - \beta$ increases.



{This figure illustrates why the replication number formula works well. It was generated with R, and it uses somewhat realistic values of the parameters. An attached document contains a proof of a related inequality. (Click on the image to magnify.) i/div