Power Analysis for a-priori determination of sample size

Martin Förster

August 2, 2022

Power Analysis

Power Analysis is actually just the conversion of an equation in which the statistical test power is a central quantity. In general, this is understood to mean the conversion according to a sample size n, which is at least necessary to achieve a certain test power. The statistical test power is the complementary probability to the β-error. The β-error describes the situation in which the null hypothesis (= "there is *no effect*") actually does not apply, but the null hypothesis is (incorrectly) accepted due to the data and the test results.

β-Error and statistical test power

In other words, we make the mistake of β-error when we interpret a statistical test in such a way that there is no effect - even though it is actually there. The smaller the effect, the greater the probability of this error: smaller effects are more likely to be overlooked than larger effects. Insofar as the statistical test power corresponds to $1 - \beta$: the greater the power, the lower the risk of overlooking an existing effect.

Power and n

In order to determine the power, *assumptions* must be made about the values of certain parameters as well as certain parameters must *be known*. A required quantity to determine the power is n. For the conversion to n a (minimum) power must be set as well as the value of the test statistic $\tilde{\theta}$ and the significance level α. The determination of $\tilde{\theta}$ is done by setting a maximum effect size, which is in turn a function of $\tilde{\theta}$ and n[1].

The power is

[1] Depending on the specific approach, sometimes the degrees of freedom *df* of the test instead of n used.

$$1-\beta = CDF_{\theta}\left(\left|\tilde{\theta}\right| - \theta_{1-\alpha}\right)$$

where CDF_{θ} is the cumulative density function of the random variable θ.

Example

Let θ a standard-normal distributed test statistic. What is needed is the minimal n to discover an effect of strength $\rho = 0.1$ at a significance level of $\alpha = 0.05$ with a power of $1 - \beta = 0.8$.

$$0.8 = CDF_{N(0,1)}\left(\left|\tilde{\theta}\right| - 1.6449\right)$$
$$0.8 = CDF\left(0.8416\right) \Rightarrow \left|\tilde{\theta}\right| = 0.8416 + 1.6449 = 2.4865$$

DASMod (Förster 2022) calculated $\rho = \sqrt{\frac{\theta^2}{\theta^2+n}}$. Accordingly

$$0.1 = \sqrt{\frac{2.4865^2}{2.4865^2 + \tilde{n}}}$$

$$\tilde{n} = \frac{\tilde{\theta}^2}{\tilde{\rho}^2} - \tilde{\theta}^2 = \frac{2.4865^2}{0.1^2} - 2.4865^2 \approx 612$$

The iterative method according to Cohen (2013) and G*Power (Faul et al. 2009), which possibly also is working with this method, both come to a similar result.

Finally, two suggestions will be given to facilitate proper interpretation. First: In this example, with $\rho = 0.1$ (in a notation according to Cohen 2013: $f^2 = 0.01$) that smallest still acceptable effect size according to conventions was chosen. In order to discover larger effects with the same power, a significantly lower $\tilde{n}$ is required. Second, $n < \tilde{n}$ does not mean that effects of strength $\rho \leq 0.1$ cannot be detected. It simply means that the probability of the β-error is at least 20%.

References

Cohen, Jacob (2013). *Statistical Power Analysis for the Behavioral Sciences.* Routledge. DOI: 10.4324/9780203771587.

Faul, Franz et al. (2009). “Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses”. In: *Behavior Research Methods* 41.4, pp. 1149–1160. DOI: 10.3758/brm.41.4.1149.

Förster, Martin (2022). *DASMod.* (2022.08.01.175). DOI: 10.17605/OSF.IO/JPFYG.