Homework 7 Problem 1

William Matzko

March 18, 2021

For this problem, I decided to test the claim that

$$n \approx \left(\frac{\sigma(z_{\alpha/2} + z_\beta)}{\mu_0 - \mu'} \right)^2$$

for a 2-tailed approximation. My code will use the above formula to calculate the approximate value of n and the corresponding $\beta(\mu')$, then calculate the true value of n such that $\beta(\mu') = \beta$. Note that the algorithm for finding the true value of n is far from perfect, but works in a fair number of cases. The code also returns the absolute error for n and $beta(\mu')$.

Running the code, we notice that the above formula for n is indeed an approximation, and a good one at that. Since we require n to be an integer, in practice meaning we take the ceiling value of n, is is very unlikely that this approximation will result in any tangible difference. Even when the ceiling value is not computed, the absolute error in provided by the approximation is quite small, usually on the order of 10^{-4} to 10^{-6}.

We can then ask why the above formula is an approximation to n, instead of an exact value. To see why this (may) be the case, we look at the Type II error probability for the two-tailed test in the box at the top of page 314. If we reverse engineer the approximation for n, we will find that we must make the assumption

$$-z_\beta = z_{\alpha/2} + \frac{\mu_0 - \mu'}{\sigma/\sqrt(n)}$$

In effect, we are missing the second term (the one involving $-z_{\alpha/2}$) in the above equation. This is where the approximation comes in. We should be subtracting an additional term here; however, that term corresponds to a small area underneath the normal distribution compared to the area given by the first term. Hence, it is generally permissible to ignore that second term.