An approach to sample size estimation, by Shelley Rasmussen

Sample Size Determination for a single sample: Suppose we want to be able to detect a difference between a sample mean and a "standard" mean ***M***, if the absolute value of the difference is at least Δ. We do not want the interval ***xbar ± c × SE(xbar)*** to contain ***M*** if the absolute value of ***(xbar – M)*** is greater than or equal to Δ. (We say such a difference is "statistically significant".) Equivalently, we do not want the interval ***(xbar – M) ± c × SE(xbar)*** to contain 0 if $|xbar - M| \geq \Delta$. Then we want ***c × SE(xbar)*** $\leq \Delta$. Since ***SE(xbar) = s/sqrt(n)***, the inequality we need to solve is: $n \geq c^2 s^2 / \Delta^2$.

Example: Our company is developing a new drug to lower blood pressure. Here is the drop in BP (mm Hg) we see in a sample of 5 volunteers: 8, 9, 11, 12, 14.

One sample t-test of mu = 10 vs not = 10

Variable	N	Mean	StDev	SE	95% CI	T	P
change in BP	5	10.80	2.39	1.07	(7.84, 13.76)	0.75	0.495

How big must the sample size be to say that the mean of the sample is "significantly" greater than 10 mm Hg? We have $\Delta = 0.8$ mm Hg. To be conservative, let **c = 2.5** and **s = 2.4**.
Then $n \geq 2.5^2 \times 2.4^2 / 0.8^2 = 56.25$. We need a sample of at least 57 volunteers. (If we let c = 2, the sample size estimate becomes n = 36.)

Another example: In the past, you have been able to "cure" 60% of patients with a given disease. You have a new treatment to test and you want to say it is better than the old treatment if it "cures" at least 70% of patients. How many patients with the disease to you need to treat with the new treatment?
We will be "conservative" and let $s^2 = 0.5 \times 0.5 = 0.25$ and let **c = 2.5**. Then $\Delta = .7 - .6 = .1$ and $n \geq 2.5^2 \times (0.5)^2 / 0.1^2 = 156.25$ or 157 patients. (If we let c = 2, the estimate is n = 100 patients.)

[Recall that to use this large sample analysis for 0/1 data, we need at least 5 observations in each of the 2 categories, which seems reasonable here.]

Sample Size Determination for comparison of 2 means: Suppose we want to be able to detect a "significant" difference between 2 population means if the absolute value of the difference is greater than or equal to Δ.
We do not want the confidence interval ***(xbar – ybar) ± c × SE(xbar-ybar)*** to contain 0 if $|xbar - ybar| \geq \Delta$ (we will say such a difference is "statistically significant").
Therefore, we want ***c × SE(xbar-ybar)*** $\leq \Delta$.
Since ***SE(xbar-ybar) = s × sqrt(2/n)***, the inequality we need to solve is: $n \geq 2c^2 s^2 / \Delta^2$.
Here s^2 is the pooled estimate of the variance of the 2 populations sampled.

Example: Maximal oxygen capacity was measured for professional women athletes in 2 sports:

basketball	speedskating (units not given)
42.3	46.1
42.9	52.0
49.6	

	N	Mean	StDev	SE Mean
basketball	3	44.93	4.05	2.3
speedskating	2	49.05	4.17	3.0

Difference = mu (basketball) - mu (speedskating) Estimate for difference: -4.12
95% CI for difference: (-16.01, 7.77)
2- Sample t-test of difference = 0 (vs not =): T-Value = -1.10 P-Value = 0.351 DF = 3
Both use Pooled StDev = 4.0927

The difference between basketball players and speedskaters is not significant for mean lung capacity. What sample size do we need to see a significant difference in average lung capacity between basketball players and speedskaters? We want to be able to detect a difference of at least 49.05-44.93=4.12 units. Let $\Delta = 4$ and $s = 4.1$ (pooled StDev shown above). To be conservative, let $c = 2.5$. Then $n \geq 2c^2 s^2 / \Delta^2$ becomes: $n \geq 2 \times 2.5^2 \times 4.1^2 / 4^2 = 13.1$. So we need to plan on 14 or more athletes in each group.