Power Analysis

Jeremy Albright

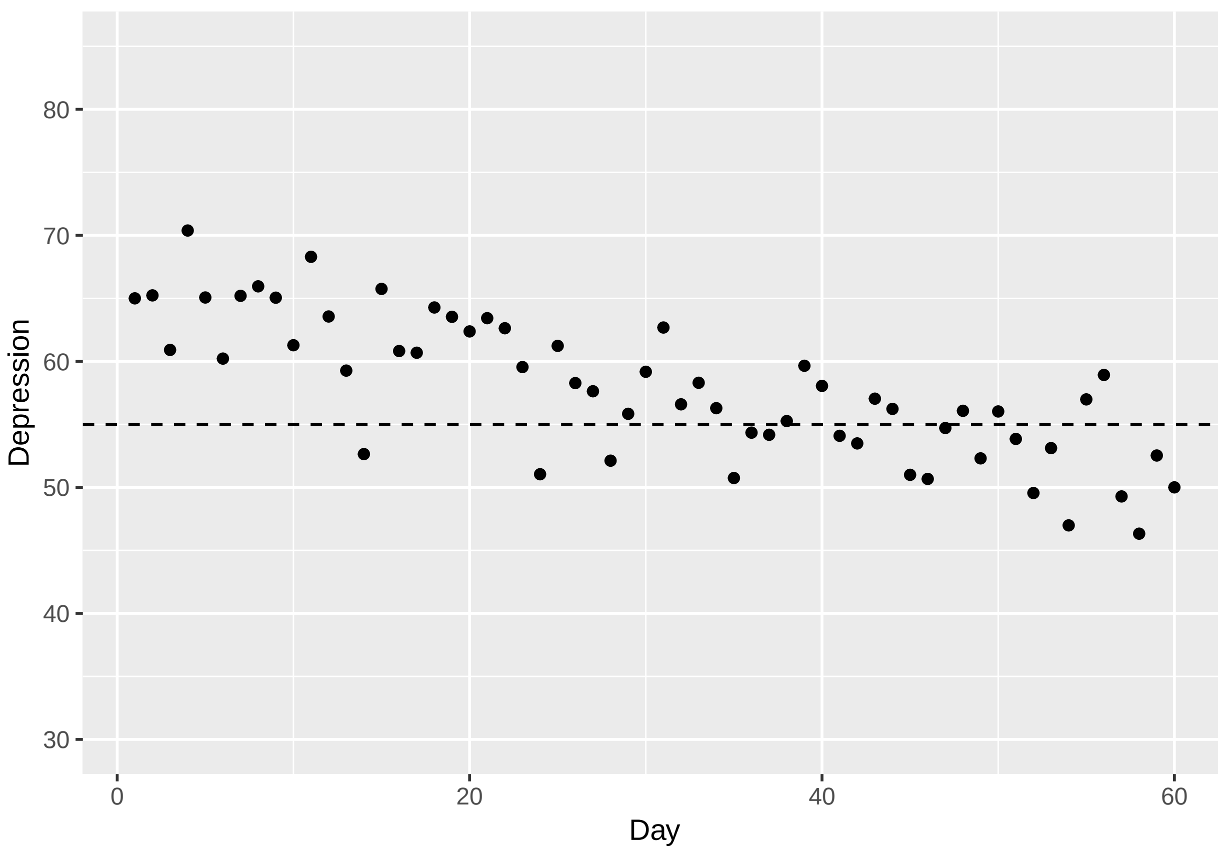
May 9, 2019

Example

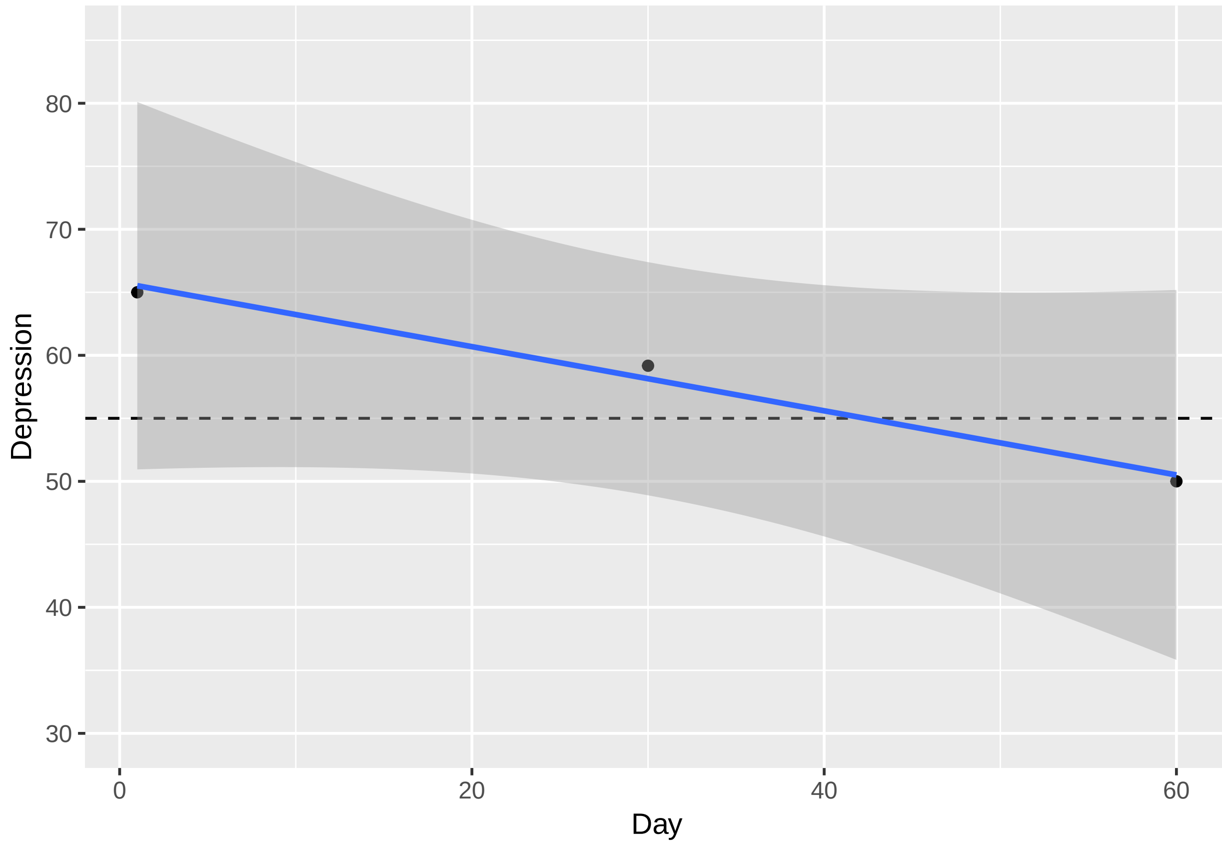
The purpose of this document is to demonstrate the improvement in power that exists from increasing the number of repeated measures. Say the goal is to improve scores on a depression scale measured as T-scores (mean of 50, standard deviation of 10), with higher values meaning higher depression. A patient begins treatment with a score of 65, or 1.5 standard deviations above the population average mood. The goal of treatment is to get the patient's score to be significantly less than 55.

Statistical Confidence as Observations Increase

Say the patient has one 30-day and one 60-day follow-up with the therapist. Her depression improves to 50 over this period, though there is some noise over time due to day-to-day variability in stress (as well as, especially for patient-reported outcomes, some measurement error). If we had data on every single day, the plot of scores would look like the following, with the horizontal line representing the treatment goal.

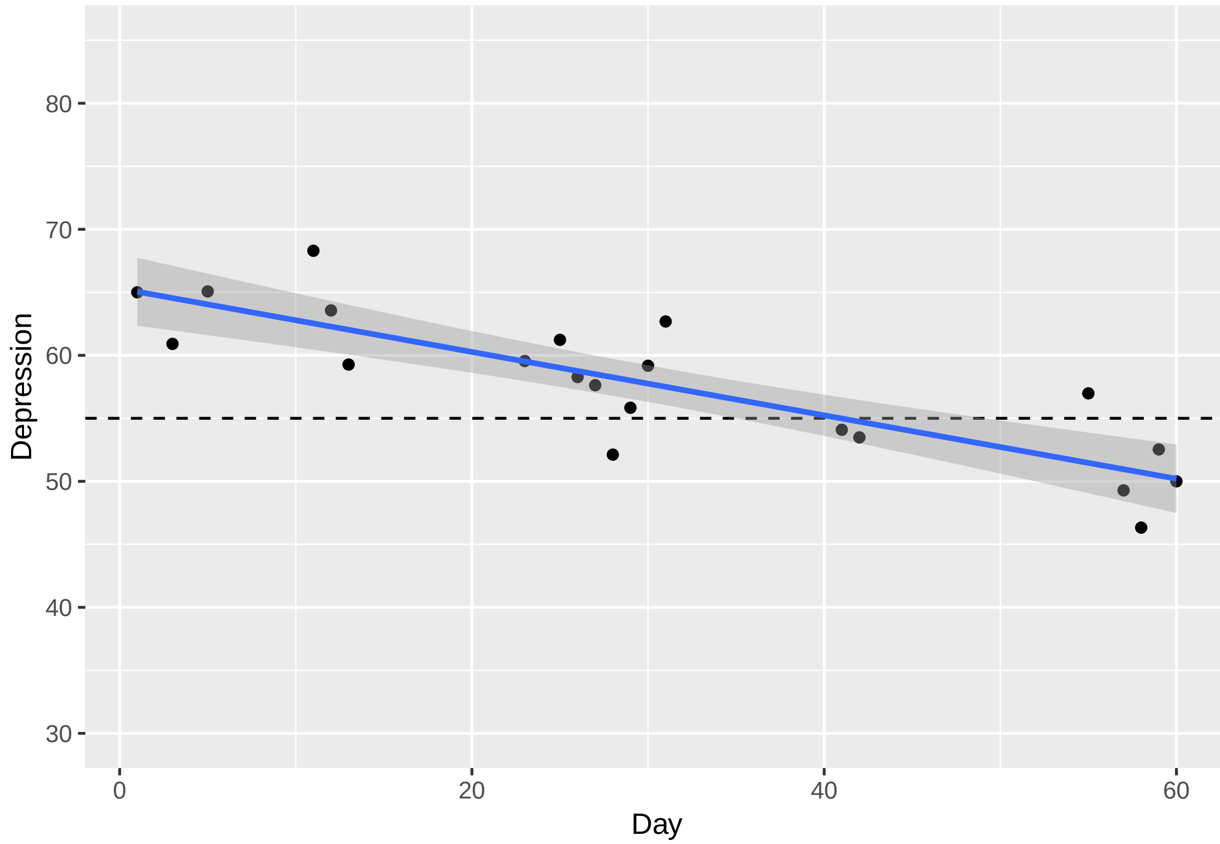


Now assume that the patient checks in with the therapist on days 30 and 60. Can we say, with 95% confidence, that the patient has met the treatment goal based on our three observations? Statistically, this translates into determining if the 95% confidence interval around the day 60 measure is entirely below 55. The following figure displays the confidence interval around the trend identified with the three data points.



Although the patient improved, the 95% confidence interval around the value at day 60 extends far above 55, meaning that one cannot rule out with 95% confidence that the patient failed to meet her goal.

However, if the patient logged her mood on roughly one-third of the days, the confidence interval looks like the following:



The 95% confidence interval at day 60 is now entirely below the treatment goal, and the therapist can have more confidence that the treatment has been effective.

Analytic Formula

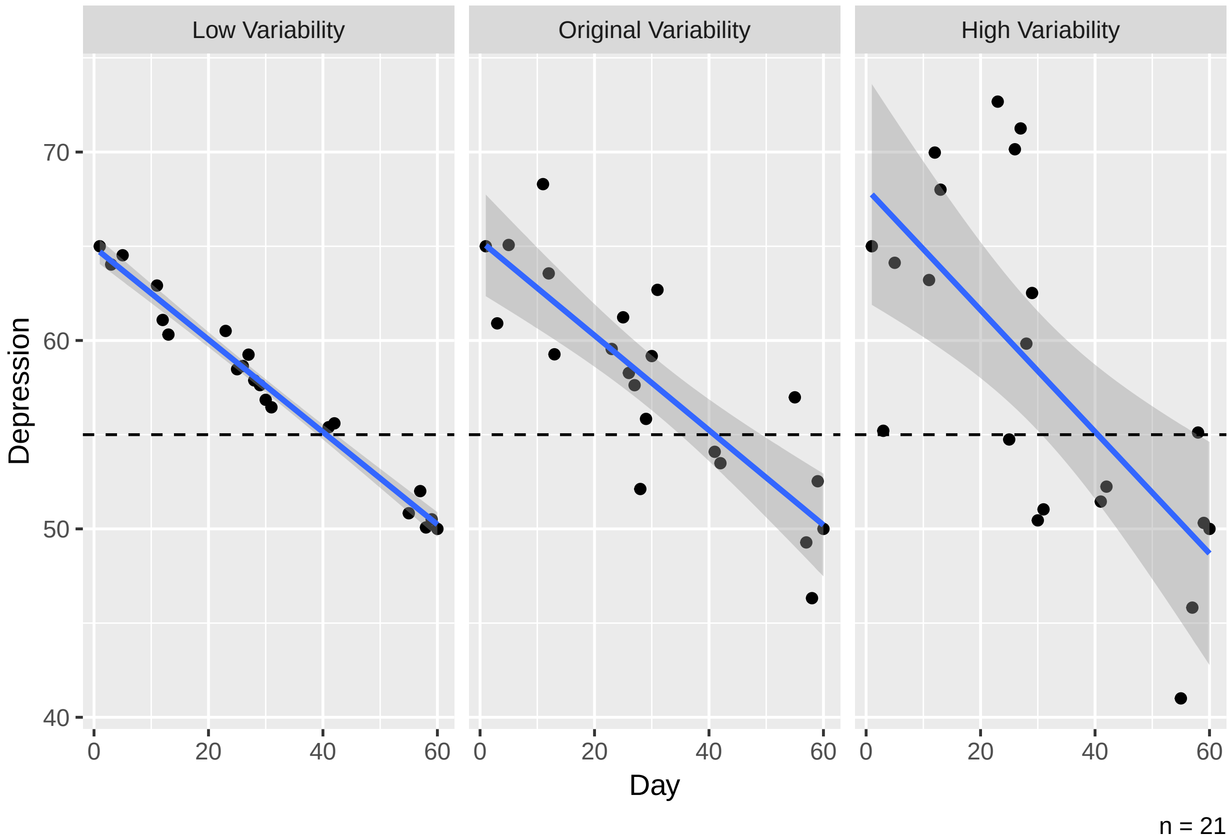
The 95% confidence interval around a predicted value $\hat{y}$ for a given value of the predictor x in a simple regression is the following:

$$\hat{y} \pm t_{n-2} s_y \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{(n-1)s_x^2}}$$

where t is the value corresponding to the 97.5th percentile from a t-distribution with $n-2$ degress of freedom, s_y is the standard deviation of the residuals (distance from each observation to the predicted line), and s_x^2 is the variance of the independent variable x (day in the above example).

One element of this formula is the sample size, n. As n only appears in the denominators, larger n leads to a smaller confidence interval and hence more precision in the estimate on a given day. Note also that s_y plays a role in the width of the confidence interval. As the amount of variability around the regression line increaes, so too does the size of the confidence interval. In other words, a measure that varies a lot from one day to the next will make it harder to have confidence in any single day's estimate.

To illustrate this latter point, we can increase and decrease the amount of variability around the regression line relative to the above example.



The upshot is that both measurement reliability *and* sample size play a role in determining the confidence with which a therapist can conclude that a treatment goal has been met.