

Factors Affecting Power

Prerequisites

[Introduction to Power](#), [Example Calculations](#), [Significance Testing](#), [Type I and Type II Errors](#), [One- and Two-Tailed Tests](#)

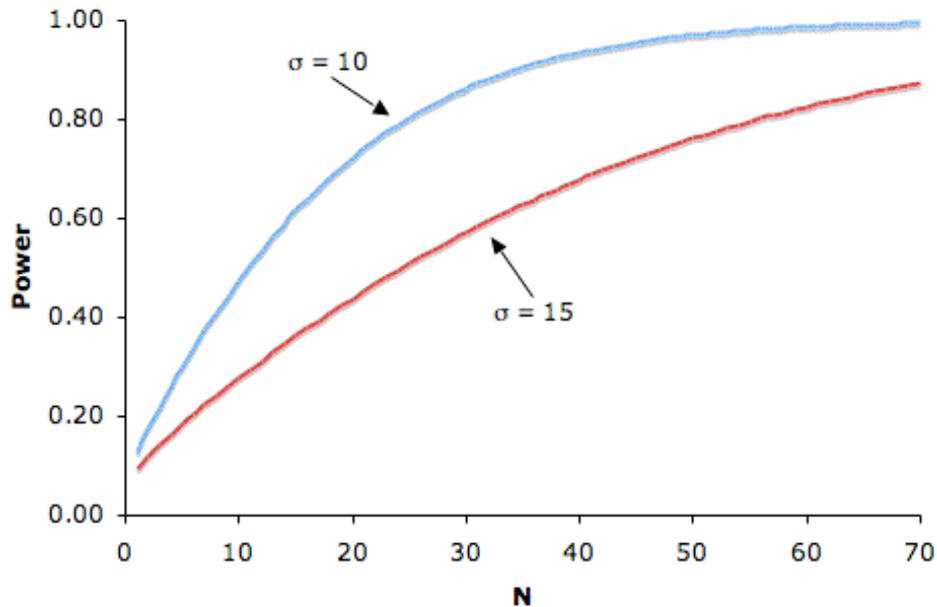
Several factors affect the power of a statistical test. Some of the factors are under the control of the experimenter whereas others are not. The following example will be used to illustrate the various factors.

Suppose a math achievement test were known to be [normally distributed](#) with a mean of 75 and [standard deviation](#) of σ . A researcher is interested in whether a new method of teaching results in a higher mean. Assume that although the experimenter does not know it, the population mean μ is larger than 75. The researcher plans to sample N subjects and do a one-tailed test of the whether the sample mean is significantly higher than 75. What is the probability that the researcher will correctly reject the false [null hypothesis](#) that the [population](#) mean is 75 at the 0.05 level?

Sample Size

Figure 1 shows that the larger the sample size, the higher the power. Since sample size is typically under an experimenter's control, increasing sample size is one way to increase power. However, it is sometimes difficult and/or expensive to use a large sample size.

Figure 1. The relationship between sample size and power for $H_0: \mu = 75$, real $\mu = 80$, one-tailed $\alpha = 0.05$, for σ 's of 10 and 15.



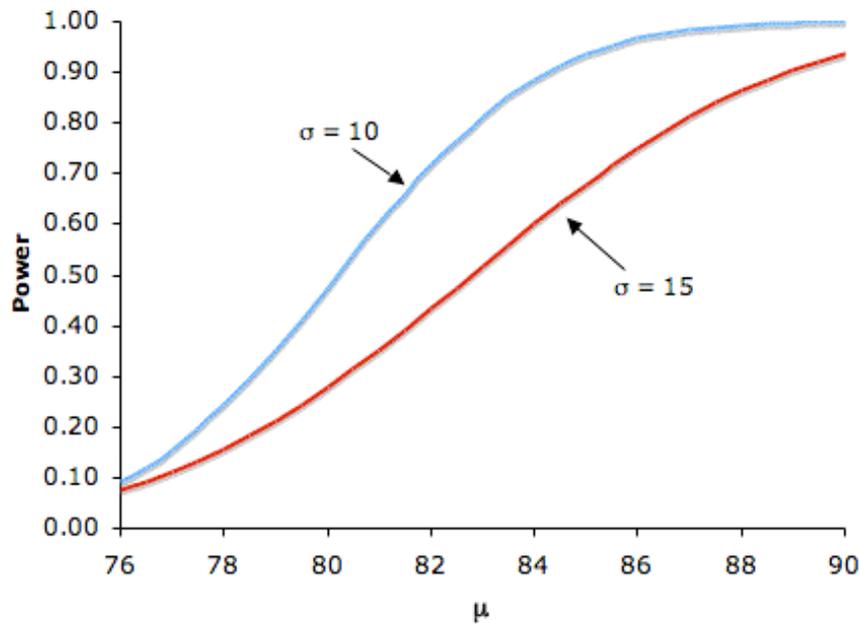
Standard Deviation

Figure 1 also shows that power is higher when the standard deviation is small than when it is large. Experimenters can sometimes control the standard deviation by sampling from a homogeneous population of subjects or by reducing random measurement error.

Difference between Hypothesized and True Mean

Naturally, the larger the effect size, the more likely it is than an experiment would find a significant effect. Figure 2 shows the effect of increasing the difference between the mean specified by the null hypothesis (75) and the population mean μ for standard deviations of 10 and 15.

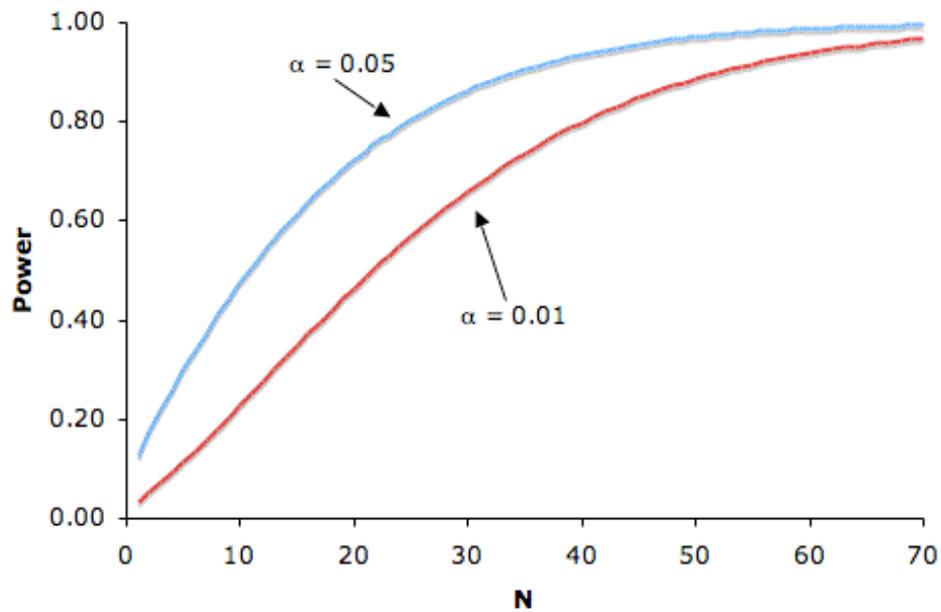
Figure 2. The relationship between power and μ with $H_0: \mu = 75$, one-tailed $\alpha = 0.05$, for σ 's of 10 and 15.



Significance Level

There is a tradeoff between the [significance level](#) and power: the more stringent (lower) the significance level, the lower the power. Figure 3 shows that power is lower for the 0.01 level than it is for the 0.05 level. Naturally, the stronger the evidence needed to reject the null hypothesis, the lower the chance that the null hypothesis will be rejected.

Figure 3. The relationship between power and significance level with one-tailed tests: $\mu = 75$, real $\mu = 80$, and $\sigma = 10$.



One- versus Two-Tailed Tests

Power is higher with a *one-tailed* test than with a *two-tailed* test as long as the hypothesized direction is correct. A one-tailed test at the 0.05 level has the same critical value as a two-tailed test at the 0.025 level. As shown the lower the significance level the lower the power. A one-tailed test, in effect, raises the significance level.