Statistical Power

In informal terms, the power of a statistical test describes the test's ability to discern a deviation from null expectations.

More Formally

Formally, power is $1 - \beta$, where β is the probability of type II error.
As you would expect, power depends upon

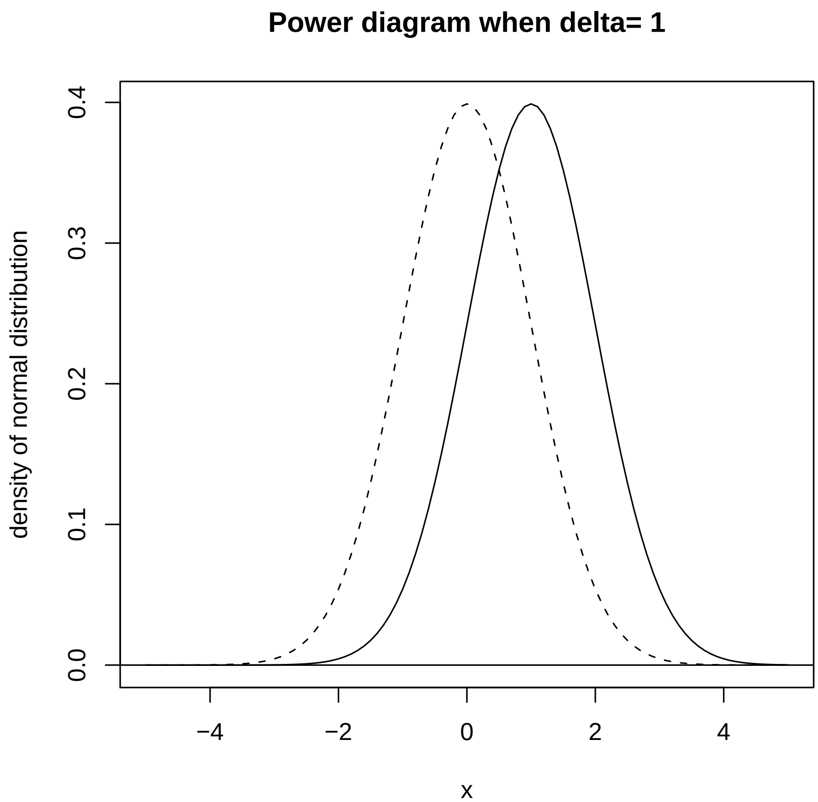
- things that influence signal
  - Differences (between means, typically)
- things that influence noise
  - sample variance
  - sample size
- desired α level

Alternative hypothesis

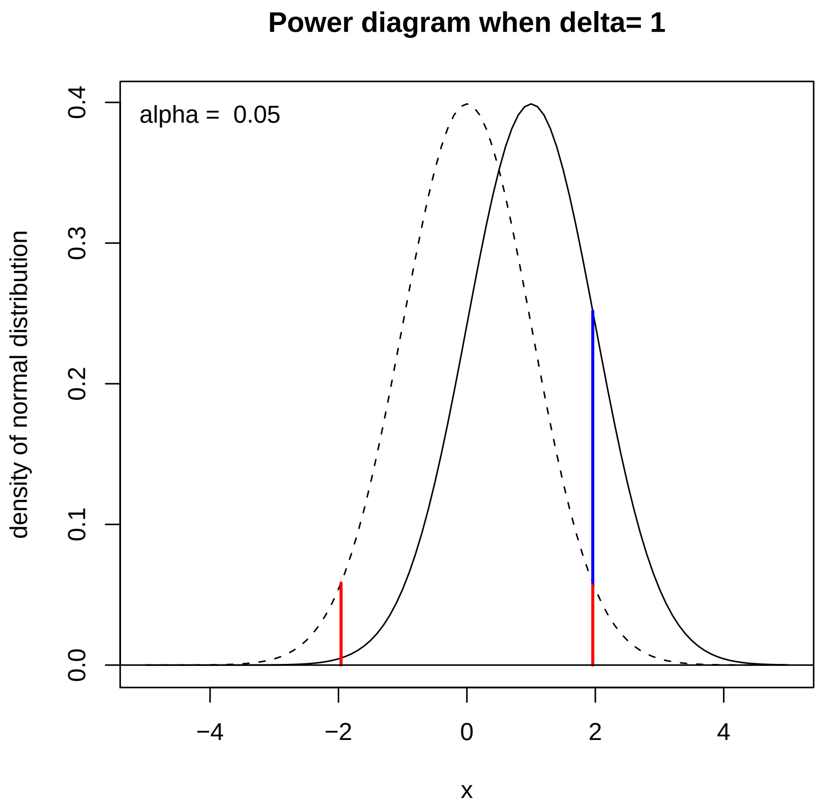
To talk about power it is necessary to define a *specific* alternative hypothesis.

- No longer:
  - $H_0 : \mu_1 = \mu_2$ and
  - $H_1 : \mu_1 \neq \mu_2$
- H_1 must be specified in such a way that the alternative is well defined in terms of the test statistic.
  - $H_0 : \mu_1 = \mu_2$ and
  - $H_1 : \mu_1 - \mu_2 = 1$
- or
  - $H_0 : \mu_1 = \mu_2$ and
  - $H_1 : \mu_1 - \mu_2 \geq 1$

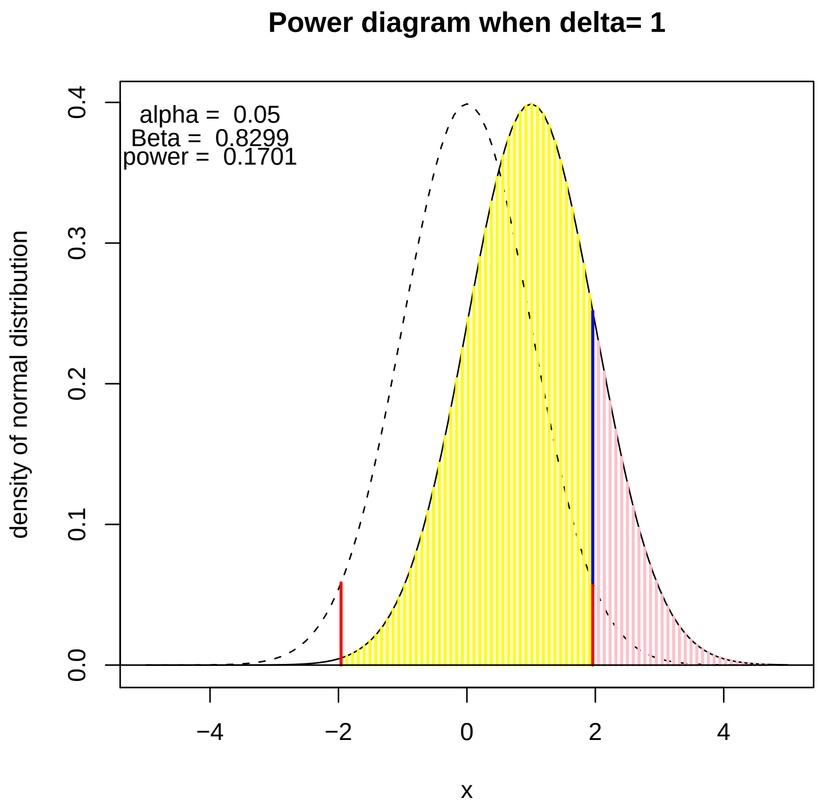
Null and alternate



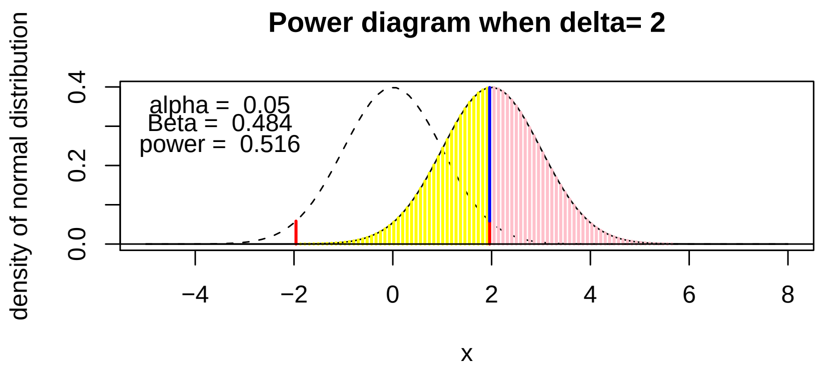
Null and alternate cont

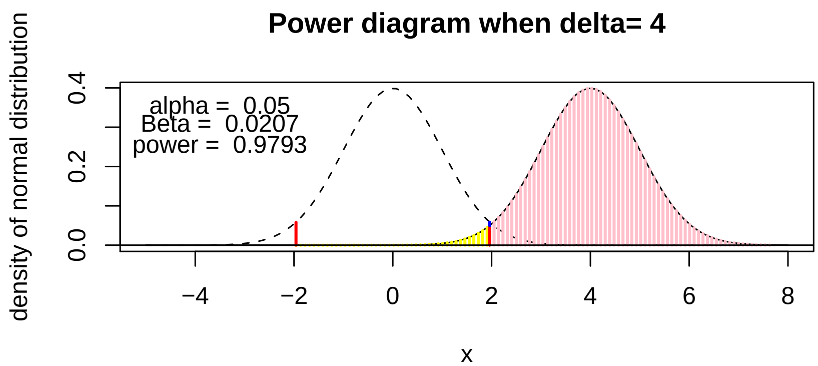


Null and alternate, cont



Different deltas





Calculating power

As you can see, when comparing draws from two normal populations with equal variance, it is relatively easy to calculate power.

- $\beta = \int_{-\infty}^{\text{critical}} N(\mu_{H_1}, \sigma_{H_1})$
- in R the `pnorm()` function does the trick
- power $= 1 - \beta$

Calculating power, cont

This analysis assumes:

- two populations
- knowledge of parametric values

More complicated power analyses have been developed for situations where the variance and means need to be estimated for two or more samples. R implements some of these calculators

```
> power.t.test(n = 10, delta = 1, sd = 1, sig.level = 0.05, power = NULL)

     Two-sample t test power calculation

              n = 10
          delta = 1
             sd = 1
      sig.level = 0.05
          power = 0.5619846
    alternative = two.sided

 NOTE: n is number in *each* group
```

two sample cont

```
> power.t.test(n = NULL, delta = 1, sd = 1, sig.level = 0.05, power = 0.95)

     Two-sample t test power calculation

              n = 26.98922
          delta = 1
             sd = 1
      sig.level = 0.05
          power = 0.95
    alternative = two.sided

 NOTE: n is number in *each* group
```

two sample cont

```
> power.t.test(n = 10, delta = NULL, sd = 1, sig.level = 0.05,
+     power = 0.95)

     Two-sample t test power calculation

              n = 10
          delta = 1.706224
             sd = 1
      sig.level = 0.05
```

```
        power = 0.95
  alternative = two.sided

NOTE: n is number in *each* group
```

When to use power analyses

Best After running a pilot study and estimating variances, use power analysis to determine sample size of actual experiment.

This is very important in many cases because it can really save money and in many cases, time (even if you have to run two experiments).

You do need to make a prediction on what you think the biologically significant treatment effects will be, but that can also be estimated from preliminary experiments in some cases

Not as good but not bad After failing to reject H_0, determine the power to reject it *post hoc*.