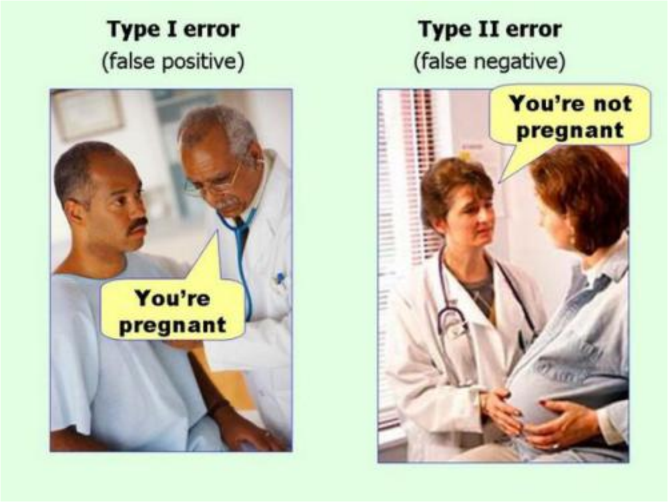
Sample Size & Power

Dr. Yen-Yi Ho

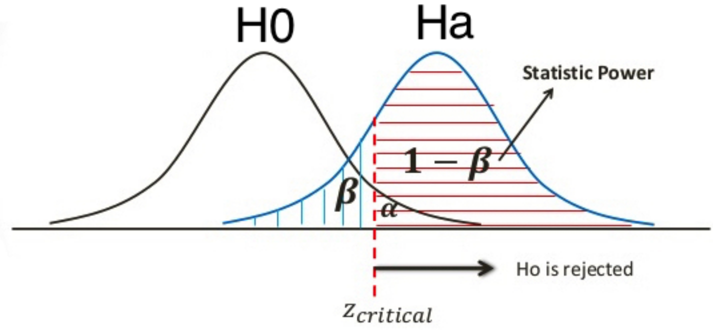
Department of Statistics, University of South Carolina

Making Mistakes



Power

- Power is the probability of rejecting the null hypothesis when it is false
- Hence, **power** (as its name would suggest) is a good thing; you want more power
- A type II error (a bad thing, as its name would suggest) is failing to reject the null hypothesis when it's false; the probability of a type II error is usually called β
- Note power= $1 - \beta$


H0
Ha
Statistic Power
$1-\beta$
β
α
Ho is rejected
$z_{critical}$

Example

- A respiratory disturbance index (RDI) of more than 30 is considered evidence of severe sleep disordered breathing.
- Suppose that in a sample of 100 participants with other risk factors for sleep disordered breathing at a sleep clinic, the mean RDI was 32 with a standard deviation of 10.

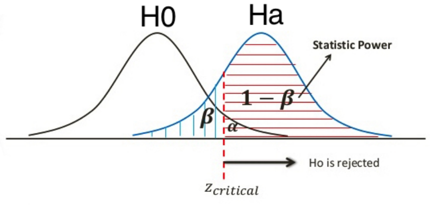
Statistical Power

- We want to test the hypothesis that
  - $H_0 : \mu = 30$
  - $H_a : \mu > 30$

  where μ is t he population mean RDI.
- The power is

$$P(\frac{\overline{X} - 30}{s/\sqrt{n}} > t_{n-1}(1-\alpha)|\mu = \mu_a)$$

- Note that this is a function that depends on the specific value of μ_a!
- Notice as μ_a approaches 30, the power approaches α



Calculating power

Assume that n is large and that we know σ

$$
\begin{aligned}
1-\beta &= P\left(\frac{\bar{X}-30}{\sigma/\sqrt{n}} > z_{1-\alpha} \mid \mu = \mu_a\right) \\
&= P\left(\frac{\bar{X}-\mu_a+\mu_a-30}{\sigma/\sqrt{n}} > z_{1-\alpha} \mid \mu = \mu_a\right) \\
&= P\left(\frac{\bar{X}-\mu_a}{\sigma/\sqrt{n}} > z_{1-\alpha} - \frac{\mu_a-30}{\sigma/\sqrt{n}} \mid \mu = \mu_a\right) \\
&= P\left(Z > z_{1-\alpha} - \frac{\mu_a-30}{\sigma/\sqrt{n}} \mid \mu = \mu_a\right)
\end{aligned}
$$

Example continued

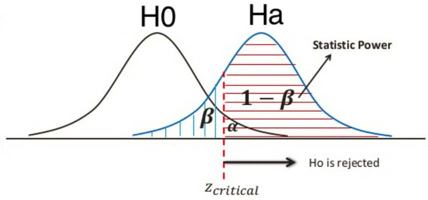
- Suppose that we wanted to detect a increase in mean RDI of at least 2 (above 30). Assume normality and that the sample in question will have a standard deviation of 4; what would be the power if we took a sample size of 16?
- $Z_{1-\alpha} = 1.645$ and $\frac{\mu_a - 30}{\sigma/\sqrt{n}} = 2/(4/\sqrt{16}) = 2$
- $P(Z > 1.645 - 2) = P(Z > -0.355) = 64\%$

Example continued

- What n would be required to get a power of 80%
- i.e. we want

$$0.80 = P\left(Z > z_{1-\alpha} - \frac{\mu_a - 30}{\sigma/\sqrt{n}} \mid \mu = \mu_a\right)$$

- Set $z_{1-\alpha} - \frac{\mu_a - 30}{\sigma/\sqrt{n}} = z_{0.20}$ and solve for n

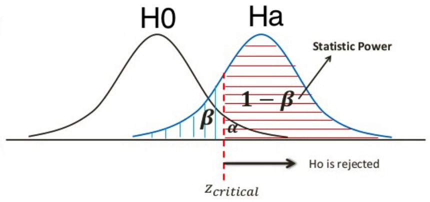


Notes

- The calculation for $H_a : \mu < \mu_0$ is similar
- For $H_a : \mu \neq \mu_0$ calculate the one sided power using $\alpha/2$ (this is only approximately right, it excludes the probability of getting a large TS in the opposite direction of the truth)

Notes

- Power goes up as α gets larger
- Power of a one sided test is greater than the power of the associated two sided test
- Power goes up as μ_1 gets further away from μ_0
- Power goes up as *n* goes up



Example

Let's recalculate power for the previous example using the T distribution instead of the normal; here's the easy way to do it. Let $\sigma = 4$ and $\mu_a - \mu_0 = 2$

```
##the easy way
power.t.test(n = 16, delta = 2 / 4,
             type = "one.sample",
             alt = "one.sided")
##result is 60%
```

Example: Monte Carlo Simulation

```
nosim<-10000
mu0<-30
mua<-32
sd<-4
n<-16
alt<-"greater"
SimData<-function(mua, sd, n){
        data<-rnorm(n=n, mean=mua, sd=sd)
        return(data)
}
Calpvalue<-function(data, mu0, alt){
        pvalue<-t.test(data, mu=mu0, alt=alt)$p.value
        return(pvalue)
}
pvalues<-rep(NA, length=nosim)
for(i in 1:nosim){
        data<-SimData(mua=mua, sd=sd, n=n)
        pvalues[i]<-Calpvalue(data, mu0=mu0, alt=alt)
}
power<-mean(pvalues<0.05)
power
##result is 60%
```

Comments

- Notice that in both cases, power required a true mean and a true standard deviation
- However in this (and most linear models) the power depends only on the mean (or change in means) divided by the standard deviation (ie: effect size $\frac{\mu_a - \mu_0}{\sigma}$)