Power

Statistical Inference

Brian Caffo, Jeff Leek, Roger Peng
Johns Hopkins Bloomberg School of Public Health

Power

- Power is the probability of rejecting the null hypothesis when it is false
- Ergo, power (as it's 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 $= 1 - \beta$

Notes

- Consider our previous example involving RDI
- $H_0 : \mu = 30$ versus $H_a : \mu > 30$
- Then power is

$$P\left(\frac{\bar X - 30}{s/\sqrt{n}} > t_{1-\alpha,n-1} \mid \mu = \boxed{\mu_a}\right)$$

also H1 ist true

- 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 for Gaussian data

Assume that n is large and that we know σ

Das ist keine Z-Statistik, weil wir 30 abziehen, und 30 ist der Mittelwert unter H0, nicht unter H1 (wo ja mu_a gilt)

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

hier mit norm.Vertl statt T-Test >> z, sigma

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

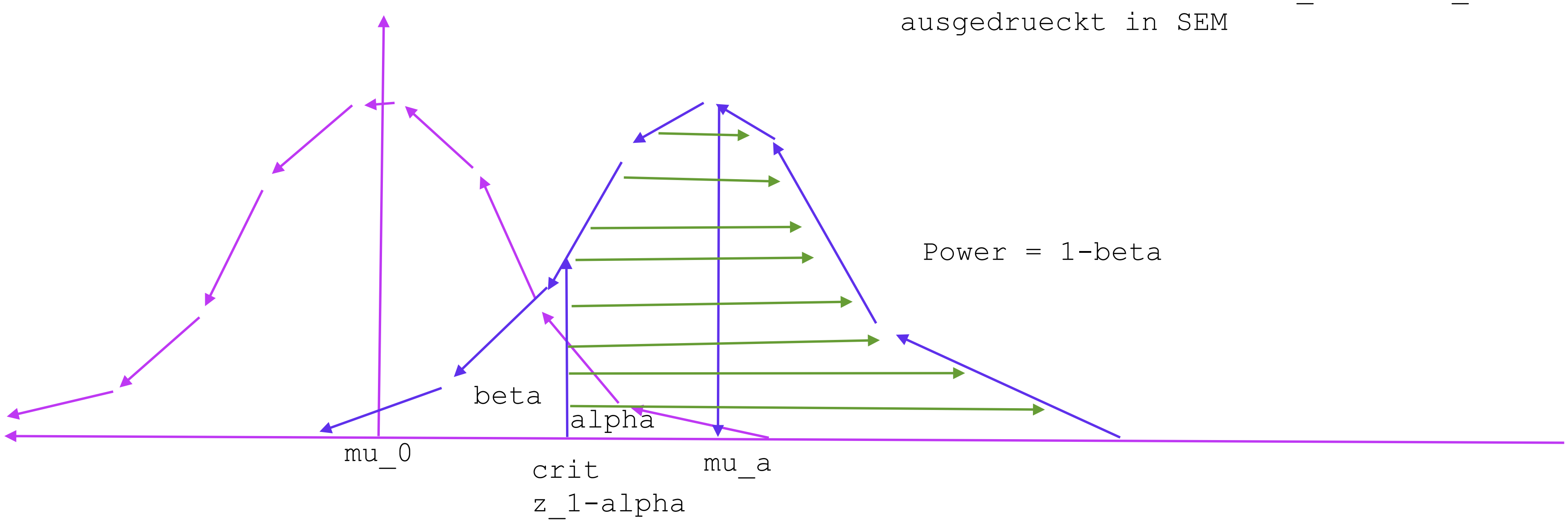
Trick: mu_a abziehen und dazuzaehlen.

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

Jetzt ist es eine Z-Statistik, weil wir mu_a abziehen, unter H1.

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

Das ist Differenz zw mu_a und mu_0 ausgedrueckt in SEM



Example continued

- Suppose that we wanted to detect a increase in mean RDI of at least 2 events / hour (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?
  - $Z_{1-\alpha} = 1.645$
  - $\frac{\mu_a - 30}{\sigma/\sqrt{n}} = 2/(4/\sqrt{16}) = 2$
  - $P(Z > 1.645 - 2) = P(Z > -0.355) = 64\%$

```
(X_ - mu_0)
─────────        ~ N(0,1) unter H0.
sigma/sqrt(n)

Unter H1 aendert sich mu und sigma:
      mu_a - mu_0
~ N(─────────────,  1)
      sigma/sqrt(0)

Daher geht auch:

pnorm(1.645, mu=2, sd=1)
```

```r
pnorm(-0.355, lower.tail = FALSE)
```

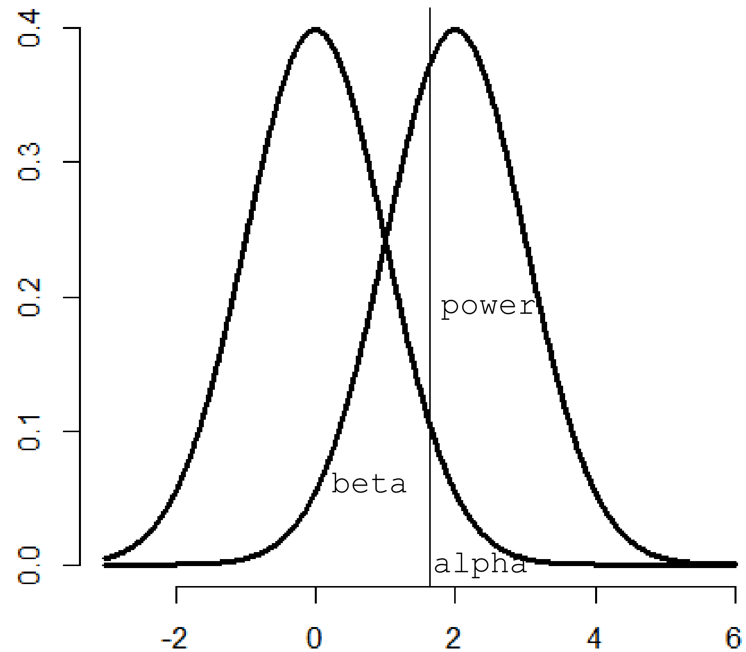
```
[1] 0.6387
```

```
Unter H0: (Xmean - mu_0)/ (sigma/sqrt(n))  ~ N(0,1)
Unter H1? immer noch ~N, mit mu = mu_a - mu_0 und var 1.
```

Note

- Consider $H_0 : \mu = \mu_0$ and $H_a : \mu > \mu_0$ with $\mu = \mu_a$ under H_a.
- Under H_0 the statistic $Z = \frac{\sqrt{n}(\bar X - \mu_0)}{\sigma}$ is $N(0,1)$
- Under H_a Z is $N\left(\frac{\sqrt{n}(\mu_a - \mu_0)}{\sigma}, 1\right)$
- We reject if $Z > Z_{1-\alpha}$

```
sigma <- 10; mu_0 = 0; mu_a = 2; n <- 100; alpha = .05
plot(c(-3, 6),c(0, dnorm(0)), type = "n", frame = false, xlab = "Z value", ylab = "")
xvals <- seq(-3, 6, length = 1000)
lines(xvals, dnorm(xvals), type = "l", lwd = 3)
lines(xvals, dnorm(xvals, mean = sqrt(n) * (mu_a - mu_0) / sigma), lwd =3)
abline(v = qnorm(1 - alpha))
```


0.4
0.3
0.2
0.1
0.0
power
beta
alpha
-2
0
2
4
6

Question

- When testing $H_a : \mu > \mu_0$, notice if power is $1 - \beta$, then

$$1 - \beta = P\left(\boxed{Z > z_{1-\alpha} - \frac{\mu_a - \mu_0}{\sigma/\sqrt{n}}} \mid \mu = \mu_a\right) = P(\boxed{Z > z_\beta})$$

- This yields the equation

$$z_{1-\alpha} - \frac{\sqrt{n}(\mu_a - \mu_0)}{\sigma} = z_\beta$$

- Unknowns: μ_a, σ, n, β
- Knowns: μ_0, α
- Specify any 3 of the unknowns and you can solve for the tail remainder

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)
- 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
- Power doesn't need μ_a, σ and n, instead only $\frac{\sqrt{n}(\mu_a - \mu_0)}{\sigma}$
  - The quantity $\boxed{\frac{\mu_a - \mu_0}{\sigma}}$ is called the effect size, the difference in the means in standard deviation units.
  - Being unit free, it has some hope of interpretability across settings

T-test power

- Consider calculating power for a Gossett's T test for our example
- The power is

$$P\left(\frac{\bar X - \mu_0}{S/\sqrt{n}} > t_{1-\alpha, n-1} ~|~ \mu = \mu_a\right)$$

- Calcuting this requires the non-central t distribution.
- `power.t.test` does this very well
  - Omit one of the arguments and it solves for it

Example

```
power.t.test(n = 16, delta = 2 / 4, sd=1, type = "one.sample",  alt = "one.sided")$power
```

```
                     delta: difference            instead of
                       in the means               2-group t-test
```

```
[1] 0.604
```

```
power.t.test(n = 16, delta = 2, sd=4, type = "one.sample",  alt = "one.sided")$power
```

```
[1] 0.604
```

```
power.t.test(n = 16, delta = 100, sd=200, type = "one.sample", alt = "one.sided")$power
```

```
[1] 0.604
```

```
>> Resultat immer gleich, da Power abhaengt von (mu0-mu1)/sigma = delta/sigma [hier = sd]
```

Example

Man kann der Fkt power.t.test immer drei Variablen geben, und sie spukt die vierte aus:

```
power.t.test(power = .8, delta = 2 / 4, sd=1, type = "one.sample",  alt = "one.sided")$n
```

```
[1] 26.14
```

```
power.t.test(power = .8, delta = 2, sd=4, type = "one.sample",  alt = "one.sided")$n
```

```
[1] 26.14
```

```
power.t.test(power = .8, delta = 100, sd=200, type = "one.sample", alt = "one.sided")$n
```

```
[1] 26.14
```

Achtung: viele ‘dials’ -> viele Moeglichkeiten, sich zu irren (oder selbst zu beluegen). Daher die Powerberechnung so simple wie moeglich machen!