HarvardX: HSPH-HMS214x Fundamentals of Clinical Trials

Sample Size and Power I

This note set is meant to supplement the material covered in the video lecture.

Null and Alternative Hypotheses

For superiority trials (RCTs), the study hypothesis is stated as a null hypothesis of no difference in the distribution of the primary endpoint between the study groups. We test the null hypothesis against a two-sided alternative.

<u>Example – CORONARY TRIAL</u>

The short-term null hypothesis was:

H_0: Patients receiving on-pump and off-pump coronary artery bypass surgery will have identical event rates at 30 days post-randomization.

H_A: The frequency of events at 30 days will differ in the two treatment groups.

Type I and Type II Errors

When we test the null hypothesis, there are two possible states of nature and two decisions:

	Truth About Risk Difference	
Test Result	H_0 True	H_A True
Reject H_0	**Type I Error**	No Error
Fail to Reject H_0	No Error	**Type II Error**

We will design a test that has a small probability of a Type I error, usually 0.05.

The power of the study is the probability that we will reject the null hypothesis when the alternative hypothesis is actually true. We would like this probability to be large, usually 0.80.

Test Statistic

To test the null hypothesis, we calculate a test statistic, T, and a critical value, C, and reject the null hypothesis if $|T| > C,$ that is, if $T > C$ or $T < -C$. To calculate power or sample size, we will focus on the significance in one direction, $T < -C$.

Example – CORONARY TRIAL

The 30-day outcome is a binary event, occurrence or non-occurrence of death or complications within 30 days of surgery.

p_T = Probability that an off-pump patient will have an event.
p_C = Probability that an on-pump patient will have an event.

H_0: $p_T = p_C$ (T and C are equally effective)
H_A: $p_T \neq p_C$ (T and C are not equally effective)

The observed difference in proportions is $\Delta = p_T - p_C$.

Assuming equal sample sizes in the two groups,

$$Var(\Delta) = \frac{p_T(1-p_T)}{n} + \frac{p_C(1-p_C)}{n}$$

Under the null hypothesis, $p_T = p_C$, define T as the difference between the observed proportions divided by the standard error of the difference.

$$T = \frac{p_T - p_C}{\sqrt{\frac{2\bar{p}(1-\bar{p})}{n}}} = \frac{\Delta}{\sqrt{Var(\Delta)}} = \frac{\Delta}{Stdev(\Delta)}$$

where $\bar{p}$ is the average event rate (it is common in statistics to use the bar notation to depict an average; if you are unfamiliar with notations you can refer to the reference in the Appendix for a quick brush-up).

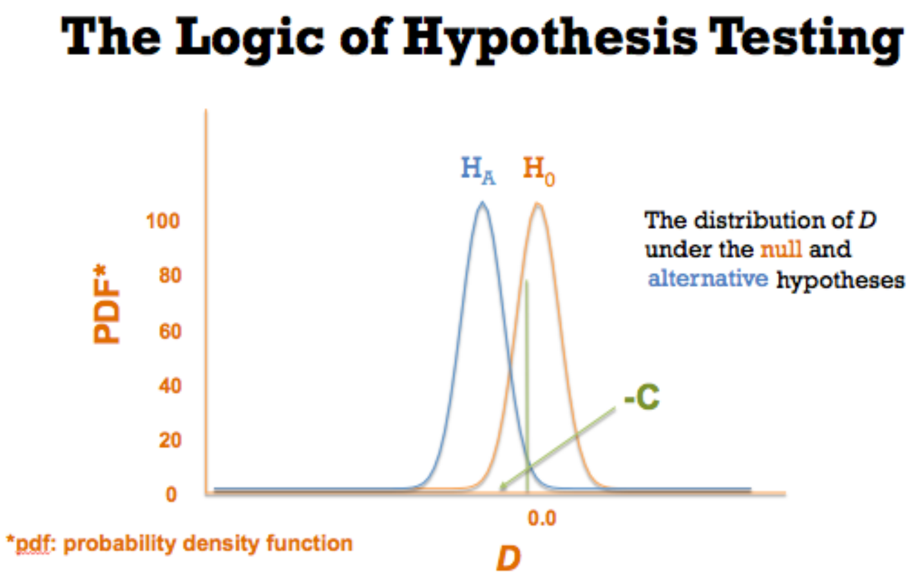
Critical Value

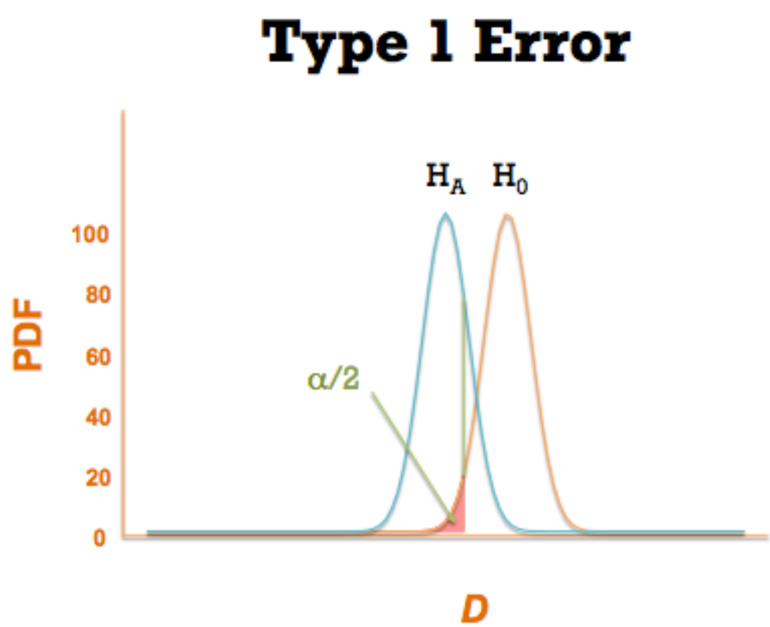
Choose the critical value C so that $P(T < -C|H_0) = \alpha/2,$ usually $\alpha = 0.05$ (two-sided) so that $\alpha/2 = 0.025$.

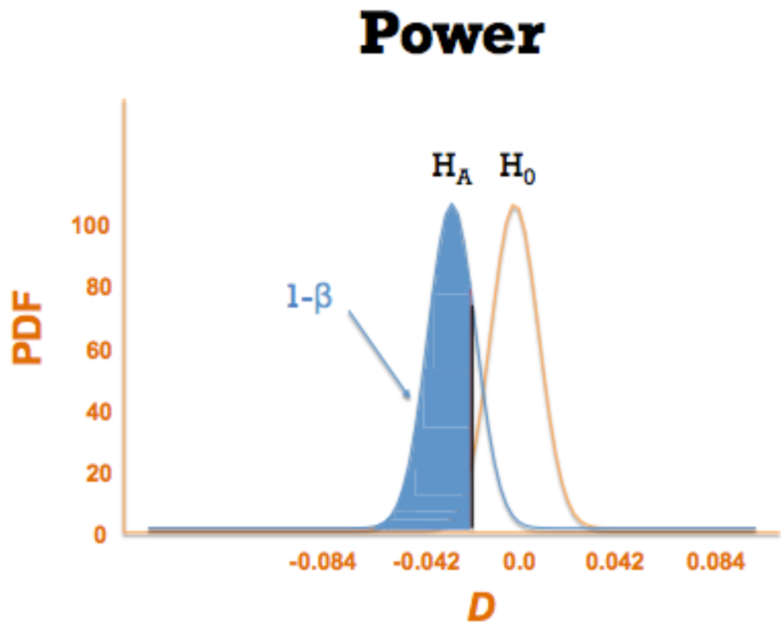
T has an approximately standard normal distribution, N(0,1), if H_0 is true. Hence, if $\alpha/2 = 0.025, C = 1.96$.

Power is $P(T < -C) given\ H_A = P(T < -C|H_A) = 1 - P(Type\ II\ Error) = 1 - \beta$

The investigator can control the power by choosing the sample size.







Power is also equal to 1- Pr(Type II Error).

Example – CORONARY TRIAL

One possible scenario for the 30-day endpoint was a 15% reduction in the event rate in the off-pump group.

$p_C = 0.08$
$p_T = 0.08 * 0.85 = 0.068$

If H_0 is true $(p_C = p_T)$, Δ is approximately normally distributed with expected value (mean) of Δ = 0 and,

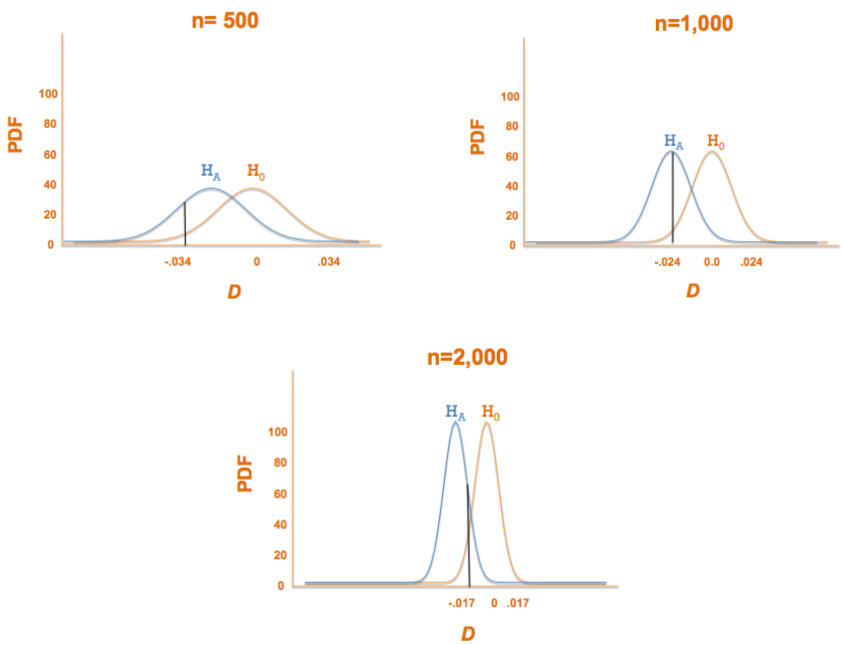
$$Var(\Delta) = \frac{2 * 0.08(1 - 0.08)}{n}$$

If H_A is true $(p_C \neq p_T)$, Δ is approximately normally distributed with expected value (mean) of $\Delta = 0.068 - 0.08 = -0.012$ and,

$$Var(\Delta) = \frac{0.068(1 - 0.068)}{n_T} + \frac{0.08(1 - 0.08)}{n_C}$$

The mean of Δ is independent of n but the variance of Δ decreases as n increases.

n	Variance of Δ	Standard dev of Δ
500	0.00029	0.017
1,000	0.00015	0.012
2,000	0.000074	0.0086



Sample Size Formula

To achieve the desired Type I and Type II errors, we use the following formula:

$$n = \frac{2\bar{p}(1-\bar{p})\left(z_{\alpha/2} + z_{\beta}\right)^2}{\Delta^2}$$

where

- $\bar{p}$ is the average event rate under H_A
- $z_{\alpha/2}$ and z_{β} are the critical values of the normal distribution
- Δ is the true difference under H_A

<u>Try it yourself:</u>

For the CORONARY Trial with $\alpha = 0.05, \beta = 0.20$, what is the appropriate sample size for each arm? (Solution on the next page)

Solution:

$p_C = 0.08$
$p_T = 0.08 * 0.85 = 0.068$
$\Delta = -0.012$
$\bar{p} = \dfrac{0.08 + 0.068}{2} = 0.074$
$z_{\alpha/2} = 1.96$
$z_\beta = 0.84$

Plug them all into this equation,

$$n = \frac{2\bar{p}(1-\bar{p})\left(z_{\alpha/2}+z_\beta\right)^2}{\Delta^2} = 7{,}461.5 \approx 7{,}462 \text{ for each arm}$$

Appendix

http://www.statistics.com/statistical-symbols/

- A useful guide to common symbols and notations used in statistics.