Power and Sample Size:
Issues in Study Design

John McGready

Department of Biostatistics, Bloomberg School

Lecture Topics

- Re-visit concept of statistical power
- Factors influencing power
- Sample size determination when comparing two means
- Sample size determination when comparing two proportions

2

Section A

Power and Its Influences

Example

- Consider the following results from a study done on 29 women, all 35–39 years old

Sample Data			
	n	Mean SBP	SD of SBP
OC users	8	132.8	15.3
Non-OC Users	21	127.4	18.2

4

Example

- Of particular interest is whether OC use is associated with higher blood pressure
- Statistically speaking we are interested in testing :

$H_o: \mu_{OC} = \mu_{NO\ OC}$ $H_o: \mu_{OC} - \mu_{NO\ OC} = 0$

$H_A: \mu_{OC} \neq \mu_{NO\ OC}$ $H_A: \mu_{OC} - \mu_{NO\ OC} \neq 0$

Here μ_{OC} represents (population) mean SBP for OC users, $\mu_{NO\ OC}$ (population) mean BP for women not using OC

5

Example

- Let's unveil the results of this test from our data:

Sample Data			
	n	Mean SBP	SD of SBP
OC users	8	132.8	15.3
Non-OC Users	21	127.4	18.2

6

Example

- The sample mean difference in blood pressures is 132.8 – 127.4 = 5.4
- This could be considered scientifically significant, however, the result is not statistically significant (or even close to it!) at the α = .05 level
- Suppose, as a researcher, you were concerned about detecting a population difference of this magnitude if it truly existed
- This particular study of 29 women has low power to detect a difference of such magnitude

Power

- Recall, from lecture four of SR1:

TRUTH

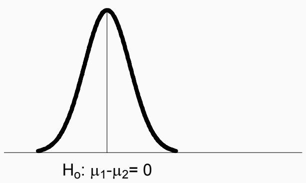
DECISION	H_o	H_A
Reject H_o	Type I Error alpha-level	Power 1-beta
Not Reject H_o		Type II Error beta

Power

- Power is a measure of “doing the right thing” when H_A is true!
- Higher power is better (the closer the power is to 1.0 or 100%)
- We can calculate power for a given study if we specify a specific H_A
- This OC/Blood pressure study has power of .13 to detect a difference in blood pressure of 5.4 or more, if this difference truly exists in the population of women 35-39 years old!
- When power is this low, it is difficult to determine whether there is no statistical difference in population means or we just could not detect it

Power

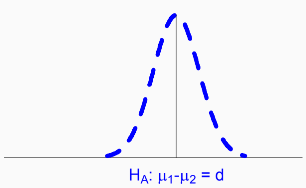
- Recall, the sampling behavior of $\bar{x}_1 - \bar{x}_2$ is normally distributed (large samples) with this sampling distributed centered at true mean difference. If H_o truth, then curve is centered at 0



10

Power

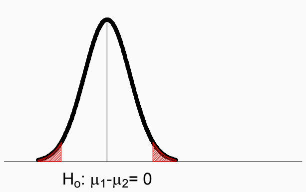
- Recall, the sampling behavior of $\bar{x}_1 - \bar{x}_2$ is normally distributed (large samples) with this sampling distributed centered at true mean difference. If H_A truth, then curve is centered at some value d, $d \neq 0$



11

Power

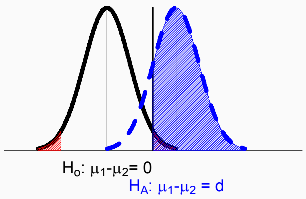
- H_o will be rejected (for α=.05) if the sample result, $\bar{x}_1 - \bar{x}_2$, is more than 2 standard errors away from 0, either above or below



12

Power

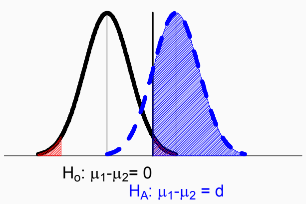
- Ho will be rejected (for α=.05) if the sample result, $\bar{x}_1 - \bar{x}_2$, is more than 2 standard errors away from 0, either above or below



13

What Influences Power?

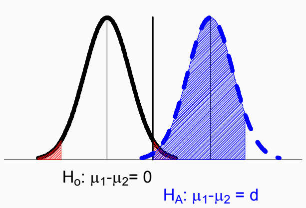
- In order to INCREASE power for a study comparing two group means, we need to do the following:
  1. Change the estimates of μ_1 and μ_2 so that the difference between the two (μ_1 - μ_2) is bigger



14

What Influences Power?

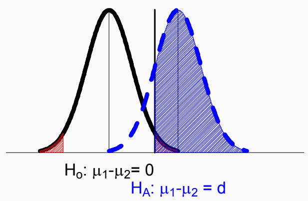
- In order to INCREASE power for a study comparing two group means, we need to do the following:
  1. Change the estimates of μ_1 and μ_2 so that the difference between the two (μ_1 - μ_2) is bigger



15

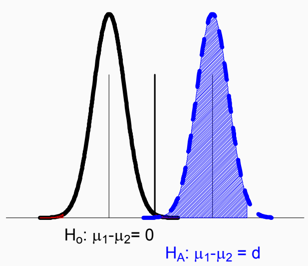
What Influences Power?

- In order to INCREASE power for a study comparing two group means, we need to do the following:
  2. Increase the sample size in each group



What Influences Power?

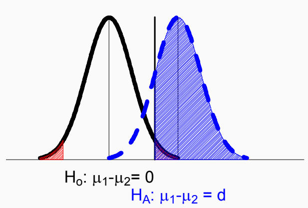
- In order to INCREASE power for a study comparing two group means, we need to do the following:
  2. Increase the sample size in each group



What Influences Power?

- In order to INCREASE power for a study comparing two group means, we need to do the following:
  3. Increase the α-level of the hypothesis test (functionally speaking, make it "easier to reject")

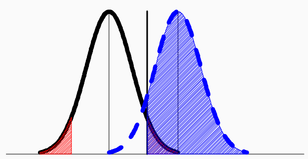
here, with α=.05:



What Influences Power?

- In order to INCREASE power for a study comparing two group means, we need to do the following:
  3. Increase the α-level of the hypothesis test (functionally speaking, make it "easier to reject")

  here, with α=.10:



H_o: μ_1-μ_2= 0
H_A: μ_1-μ_2 = d

19

Power and Studies

- Power can be computed after a study is completed
  - Can only be computed for specific H_A's: i.e. this study had XX% to detect a difference in population means of YY or greater.
  - Sometimes presented as an "excuse" for non statistically significant finding: "the lack of a statically significant association between A and B could be because of low power (< 15%) to detect a mean difference of YY or greater between.."
  - Can also be presented to corroborate with a non statistically significant result
- Many times, in study design, a required sample size is computed to actually achieve a certain preset power level to find a "Clinically/scientifically" minimal important difference in means
  - In next section we will show how to do this using Stata
  - "Industry standard" for power: 80% (or greater)

20

Section B

Sample Size Calculations when Comparing Group Means

Example

- Blood pressure and oral contraceptives
  - Suppose we used data from the example in Section A to motivate the following question:
  - Is oral contraceptive use associated with higher blood pressure among individuals between the ages of 35–39?

22

Pilot Study

- Recall, the data:

Sample Data

	n	Mean SBP	SD of SBP
OC users	8	132.8	15.3
Non-OC Users	21	127.4	18.2

23

Pilot Study

- We think this research has a potentially interesting association
- We want to do a bigger study
  - We want this larger study to have ample power to detect this association, should it really exist in the population
- What we want to do is determine sample sizes needed to detect about a 5mm increase in blood pressure in O.C. users with 80% power at significance level $\alpha = .05$
  - Using pilot data, we estimate that the standard deviations are 15.3 and 18.2 in O.C. and non-O.C. users respectively

24

Pilot Study

- Here we have a desired power in mind and want to find the sample sizes necessary to achieve a power of 80% to detect a population difference in blood pressure of five or more mmHg between the two groups

25

Pilot Study

- We can find the necessary sample size(s) of this study if we specify .
  - α-level of test (.05)
  - Specific values for μ_1 and μ_2 (specific H_A) and hence $d = \mu_1 - \mu_2$: usually represents the minimum scientific difference of interest)
  - Estimates of σ_1 and σ_2
  - The desired power(.80)

26

Pilot Study

- How can we specify $d = \mu_1 - \mu_2$ and estimate population SDs?
  - Researcher knowledge—experience makes for good educated guesses
  - Make use of pilot study data!

27

Pilot Study

- Fill in blanks from pilot study
  - α -level of test (.05)
  - Specific HA (μ_{OC}=132.8, $\mu_{NO\ OC}$ =127.4), and hence d= μ_1-μ_2 =5.4 mmHg
  - Estimates of σ_{OC} (= 15.3) and $\sigma_{NO\ OC}$ (=18.2)
  - The power we desire (.80)

28

Pilot Study

- Given this information, we can use Stata to do the sample size calculation
- "sampsi" command
  - Command syntax (items in italics are numbers to be supplied by researcher)

sampsi μ_1 μ_2, alpha(α) power(*power*) sd1(σ_1) sd2(σ_2)

29

Stata Results

- Blood Pressure/OC example

```
. sampsi 132.8 127.4, alpha(.05) power(.8) sd1(15.3) sd2(18.2)

Estimated sample size for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:

         alpha =   0.0500  (two-sided)
         power =   0.8000
            m1 =    132.8
            m2 =    127.4
           sd1 =     15.3
           sd2 =     18.2
         n2/n1 =     1.00

Estimated required sample sizes:

            n1 =      153
            n2 =      153
```

30

Pilot Study/Stata Results

- Our results from Stata suggest that in order to detect a difference in B.P. of 5.4 units (if it really exists in the population) with high (80%) certainty, we would need to enroll 153 O.C. users and 153 non-users
- This assumed that we wanted equal numbers of women in each group!

31

Stata Resulss

- Blood Pressure/OC example

```
. sampsi 132.8 127.4, alpha(.05) power(.8) sd1(15.3) sd2(18.2)

Estimated sample size for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:

         alpha =   0.0500  (two-sided)
         power =   0.8000
            m1 =    132.8
            m2 =    127.4
           sd1 =     15.3
           sd2 =     18.2
         n2/n1 =     1.00

Estimated required sample sizes:

            n1 =      153
            n2 =      153
```

32

Pilot Study/Stata Results

- Suppose we estimate that the prevalence of O.C. use amongst women 35–39 years of age is 20%
  - We wanted this reflected in our group sizes
- If O.C. users are 20% of the population, non-OC users are 80%
  - There are four times as many non-users as there are users (4:1 ratio)

33

Pilot Study/Stata Results

- We can specify a ratio of group sizes in Stata
  - Again, using "sampsi" command with "ratio" option

sampsi μ_1 μ_2, alpha(α) power(*power*) sd1(σ_1) sd2(σ_2) ratio(n_2/n_1)

Stata Resulss

- Blood Pressure/OC example

```
. sampsi 132.8 127.4, alpha(.05) power(.8) sd1(15.3) sd2(18.2) ratio(4)

Estimated sample size for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:

         alpha =   0.0500  (two-sided)
         power =   0.8000
            m1 =    132.8
            m2 =    127.4
           sd1 =     15.3
           sd2 =     18.2
         n2/n1 =     4.00

Estimated required sample sizes:

            n1 =       86
            n2 =      344
```


Section C

Sample Size Determination for Comparing Two Proportions

Power for Comparing Two Proportions

- Same ideas as with comparing means

37

Pilot Study

- We can find the necessary sample size(s) of this study if we specify .
  - α-level of test
  - Specific values for p_1 and p_2 (specific H_A) and hence d= p_1-p_2: usually represents the minimum scientific difference of interest)
  - The desired power

38

Pilot Study

- Given this information, we can use Stata to do the sample size calculation
- "sampsi" command
  - Command syntax (items in italics are numbers to be supplied by researcher)

sampsi p_1 p_2, alpha(α) power(*power*)

39

Another Example

- Two drugs for treatment of peptic ulcer compared (Familiari, et al., 1981)
  - The percentage of ulcers healed by pirenzepine (drug A) and trithiozine (drug B) was 77% and 58% based on 30 and 31 patients respectively (p-value = .17), 95% CI for difference in proportions healed ($p_{DRUG\,A}$ - $p_{DRUG\,B}$) was(-.04, .42)
  - The power to detect a difference as large as the sample results with samples of size 30 and 31 respectively is only 25%

	Healed	Not Healed	Total
Drug A	23	7	30
Drug B	18	13	31

Notes Available

Continued

Example

- As a clinician, you find the sample results intriguing – want to do a larger study to better quantify the difference in proportions healed
- Redesign a new trial, using aforementioned study results to "guestimate" population characteristics
  - Use $p_{DRUG\,A}$ = .77 and $p_{DRUG\,B}$ = .58
  - 80% power
  - α = .05
- Command in Stata
  sampsi .77 .58, alpha (.05) power (.8)

Example

- Command in Stata
  sampsi .77 .58, alpha (.05) power (.8)

```
. sampsi  .77 .58, alpha (.05) power (.8)

Estimated sample size for two-sample comparison of proportions

Test Ho: p1 = p2, where p1 is the proportion in population 1
                    and p2 is the proportion in population 2
Assumptions:

         alpha =   0.0500  (two-sided)
         power =   0.8000
            p1 =   0.7700
            p2 =   0.5800
         n2/n1 =   1.00

Estimated required sample sizes:

            n1 =      105
            n2 =      105
```

Example

- What would happen if you change power to 90%?
  sampsi .77 .58, alpha (.05) power (.9)

```
. sampsi  .77 .58, alpha (.05) power (.9)

Estimated sample size for two-sample comparison of proportions

Test Ho: p1 = p2, where p1 is the proportion in population 1
                    and p2 is the proportion in population 2
Assumptions:

         alpha =   0.0500  (two-sided)
         power =   0.9000
            p1 =   0.7700
            p2 =   0.5800
         n2/n1 =   1.00

Estimated required sample sizes:

            n1 =      136
            n2 =      136
```

43

Example

- Suppose you wanted two times as many people on trithiozone ("Drug B") as compared to pirenzephine ("Drug A")
  - Here, the ratio of sample size for Group 2 to Group 1 is 2.0
- Can use "ratio" option in "sampsi" command

44

Example

- Changing the ratio
  sampsi .77 .58, alpha (.05) power (.9) ratio(2)

```
. sampsi  .77 .58, alpha (.05) power (.9) ratio(2)

Estimated sample size for two-sample comparison of proportions

Test Ho: p1 = p2, where p1 is the proportion in population 1
                    and p2 is the proportion in population 2
Assumptions:

         alpha =   0.0500  (two-sided)
         power =   0.9000
            p1 =   0.7700
            p2 =   0.5800
         n2/n1 =   2.00

Estimated required sample sizes:

            n1 =      103
            n2 =      206
```

45

Sample Size for Comparing Two Proportions

- A randomized trial is being designed to determine if vitamin A supplementation can reduce the risk of breast cancer
  - The study will follow women between the ages of 45–65 for one year
  - Women were randomized between vitamin A and placebo

- What sample sizes are recommended?

46

Breast Cancer/Vitamin A Example

- Design a study to have 80% power to detect a 50% *relative reduction* in risk of breast cancer w/vitamin A

  (i.e. $RR = \frac{p_{VITA}}{p_{PLACEBO}} = .50$)

using a (two-sided) test with significance level α-level = .05

- To get estimates of proportions of interest:

  - using other studies, the breast cancer rate in the controls can be assumed to be 150/100,000 per year

47

Breast Cancer/Vitamin A Example

- A 50% relative reduction: if

$RR = \frac{p_{VITA}}{p_{PLACEBO}} = .50$ then, $p_{VITA} = .50 \times p_{PLACEBO}$

So, for this desired difference in the relative scale:

$$p_{VITA} = \frac{150}{100,000} \times 0.5 = \frac{75}{100,000}$$

Notice, that this difference on the absolute scale, $p_{VITA} - p_{PLACEBO}$, is much smaller in magnitude:

$$-\frac{75}{100,000} = -0.00075 = -0.075\%$$

48

Breast Cancer Sample Size Calculation in Stata

- "sampsi" command
  sampsi .00075 .0015, alpha(.05) power(.8)

```
. sampsi .00075 .0015, alpha(.05) power(.8)

Estimated sample size for two-sample comparison of proportions

Test Ho: p1 = p2, where p1 is the proportion in population 1
                    and p2 is the proportion in population 2
Assumptions:

         alpha =   0.0500  (two-sided)
         power =   0.8000
            p1 =   0.0008
            p2 =   0.0015
         n2/n1 =   1.00

Estimated required sample sizes:

            n1 =     33974
            n2 =     33974
```

49

Breast Cancer Sample Size Calculation in Stata

- You would need about 34,000 individuals per group
- Why so many?
  - Difference between two hypothesized proportions is very small: = .00075

We would expect about 50 cancer cases among the controls and 25 cancer cases among the vitamin A group

$$placebo: \frac{150}{100{,}000} \times 34{,}000 = 51$$

$$vitamin\ A: \frac{75}{100{,}000} \times 34{,}000 \approx 25$$

50

Breast Cancer/Vitamin A Example

- Suppose you want 80% power to detect only a 20% (relative) reduction in risk associated with vitamin A

A 20% relative reduction: if $RR = \frac{p_{VITA}}{p_{PLACEBO}} = .80$ then $p_{VITA} = .80 \times p_{PLACEBO}$

So, for this desired difference in the relative scale:

$$p_{VITA} = \frac{150}{100{,}000} \times 0.8 = \frac{120}{100{,}000}$$

Notice, that this difference on the absolute scale, $p_{VITA} - p_{PLACEBO}$, is much smaller in magnitude:

$$-\frac{30}{100{,}000} = -0.0003 = -0.03\%$$

51

Breast Cancer Sample Size Calculation in Stata

- "sampsi" command
  sampsi .0012 .0015, alpha(.05) power(.8)

```
. sampsi .0012 .0015, alpha(.05) power(.8)

Estimated sample size for two-sample comparison of proportions

Test Ho: p1 = p2, where p1 is the proportion in population 1
                    and p2 is the proportion in population 2
Assumptions:

         alpha =   0.0500  (two-sided)
         power =   0.8000
            p1 =   0.0012
            p2 =   0.0015
         n2/n1 =   1.00

Estimated required sample sizes:

            n1 =   241769
            n2 =   241769
```

52

Breast Cancer—Vitamin A Example Revisited

- You would need about 242,000 per group!
- We would expect 360 cancer cases among the placebo group and 290 among vitamin A group

53

An Alternative Approach—Design a Longer Study

- Proposal
  - Five-year follow-up instead of one year
    Here:
    $p_{VITA} \approx 5 \times .0012 = .006$
    $p_{PLACEBO} \approx 5 \times .0015 = .0075$
- Need about 48,000 per group
  - Yields about 290 cases among vitamin A and 360 cases among placebo
- Issue
  - Loss to follow-up

54

Section D

Sample Size and Study Design Principles:
A Brief Summary

Designing Your Own Study

- When designing a study, there is a tradeoff between :
  - Power
  - α-level
  - Sample size
  - Minimum detectable difference (specific Ha)
- Industry standard—80% power, α = .05

56

Designing Your Own Study

- What if sample size calculation yields group sizes that are too big (i.e., can not afford to do study) or are very difficult to recruit subjects for study?
  - Increase minimum difference of interest
  - Increase α-level
  - Decrease desired power

57

Designing Your Own Study

- Sample size calculations are an important part of study proposal
  - Study funders want to know that the researcher can detect a relationship with a high degree of certainty (should it really exist)
- Even if you anticipate confounding factors, these approaches are the best you can do and are relatively easy
- Accounting for confounders requires more information and sample size has to be done via computer simulation—consult a statistician!

58

Designing Your Own Study

- When would you calculate the power of a study?
  - Secondary data analysis
  - Data has already been collected, sample size is fixed
  - *Pilot Study*—to illustrate that low power may be a contributing factor to non-significant results and that a larger study may be appropriate

59

Designing Your Own Study

- What is this specific alternative hypothesis?
  - *Power* or *sample size* can only be calculated for a specific alternative hypothesis
  - When comparing two groups this means estimating the true population means (proportions) for each group

60

Designing Your Own Study

- What is this specific alternative hypothesis?
  - Therefore specifying a difference between the two groups
  - This difference is frequently called minimum detectable difference or effect size, referring to the minimum detectable difference with scientific interest

61

Designing Your Own Study

- Where does this specific alternative hypothesis come from?
  - Hopefully, not the statistician!
  - As this is generally a quantity of scientific interest, it is best estimated by a knowledgeable researcher or pilot study data
  - This is perhaps the most difficult component of sample size calculations, as there is no magic rule or "industry standard"

62

FYI—Using Stata to Compute Power

- I promised you in part A of this lecture that I would eventually show you how to compute the power to detect difference in a study that has already been conducted

- The "sampsi" command is still the command for this—we just need to feed it slightly different information for it to compute power

63

Calculating Power

- In order to calculate power for a study comparing two population means, we need the following:
  - Sample size for each group
  - Estimated (population) means, μ_1 and μ_2 for each group—these values frame a specific alternative hypothesis (usually minimum difference of scientific interest)
  - Estimated (population) SD's, σ_1 and σ_2
  - α-level of the hypothesis test

64

Calculating Power

- The Blood Pressure/Oral Contraceptive Example

Sample Data

	n	Mean SBP	SD of SBP
OC users	8	132.8	15.3
Non-OC Users	21	127.4	18.2

65

Calculating Power

- Fill in information below with results from this study
  - Sample size for each group (n_{OC} = 8, $n_{NO\ OC}$= 21)
  - Estimated (population) means, μ_{OC}= 132.8 and $\mu_{NO\ OC}$ = 127.4
  - Estimated (population) sd's, σ_{OC}= 15.3 and $\sigma_{NO\ OC}$ = 18.2 for each group
  - α-level of the hypothesis test (.05)

66

Calculating Power

- Using *sampsi* in Stata

```
. sampsi 132.8 127.4, sd1(15.3) sd2(18.2) n1(8) n2(21)

Estimated power for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:

         alpha =   0.0500  (two-sided)
            m1 =    132.8
            m2 =    127.4
           sd1 =     15.3
           sd2 =     18.2
sample size n1 =        8
            n2 =       21
         n2/n1 =     2.63

Estimated power:

         power =   0.1268
```

67

Calculating Power

- In order to calculate power for a study comparing two population proportions, we need :
  - Sample size for each group
  - Estimated (population) proportions, p_1 and p_2 for each group : these values frame a specific alternative hypothesis (it usually is the minimum difference of scientific interest)
  - α-level of the hypothesis test

68

Calculating Power

- Ulcer Drug/Healing Example

	Healed	Not Healed	Total
Drug A	23	7	30
Drug B	18	13	31

In this study:

$$\hat{p}_{DRUG\ A} = \frac{23}{30} \approx .77$$

$$\hat{p}_{DRUG\ B} = \frac{18}{31} \approx .58$$

69

Calculating Power

- Fill in information below with results form this study
  - Sample size for each group ($n_{DRUG\ A}$ = 30, $n_{DRUG\ B}$= 31)
  - Estimated (population) proportions, $p_{DRUG\ A}$ = .77 and $p_{DRUG\ B}$= .58
  - α-level of the hypothesis test (.05)

70

Calculating Power

- Using the sampsi command
  - sampsi p_1 p_2, n1(n_1) n2(n_2) alpha(α)

```
. sampsi .77 .58, n1(30) n2(31) alpha(.05)

Estimated power for two-sample comparison of proportions

Test Ho: p1 = p2, where p1 is the proportion in population 1
                  and p2 is the proportion in population 2
Assumptions:

         alpha =   0.0500  (two-sided)
            p1 =   0.7700
            p2 =   0.5800
sample size n1 =       30
            n2 =       31
         n2/n1 =     1.03

Estimated power:

         power =   0.2525

Note: For the above sample size(s) and proportion(s), the normal approximation to the
      binomial may not be very accurate.  Thus, power calculations are questionable.
```

71

Section E

FYI if Interested

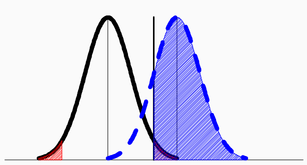
Example

- Consider the following results from a study done on 29 women, all 35-39 years old

Sample Data			
	n	Mean SBP	SD of SBP
OC users	8	132.8	15.3
Non-OC Users	21	127.4	18.2

Example

- Suppose we want to design a study with 80% power to detect a mean difference of at least 5.4 mmHg between the two group
  - will reject at α=.05 if $\left|\frac{\bar{x}_{OC}-\bar{x}_{NO\,OC}}{SE(\bar{x}_{OC}-\bar{x}_{NO\,OC})}\right| \geq 2$
  - as such, want $\Pr\left(\left|\frac{\bar{x}_{OC}-\bar{x}_{NO\,OC}}{SE(\bar{x}_{OC}-\bar{x}_{NO\,OC})}\right| \geq 2\right) = 0.80$ if $\mu_{OC}-\mu_{NO\,OC} \geq$ 5.4mmHg



Example

- Consider $SE(\bar{x}_{OC}-\bar{x}_{NO\,OC})$

Using the estimates from the small study for population SDs:

$$SE(\bar{x}_{OC}-\bar{x}_{NO\,OC}) = \sqrt{\frac{\sigma^2_{OC}}{n_{OC}} + \frac{\sigma^2_{NO\,OC}}{n_{NO\,OC}}}$$

with $n_{OC}=n_{NO\,OC}=n$ this becomes:

$$SE(\bar{x}_{OC}-\bar{x}_{NO\,OC}) = \sqrt{\frac{\sigma^2_{OC}}{n} + \frac{\sigma^2_{NO\,OC}}{n}} = \sqrt{\frac{1}{n}(\sigma^2_{OC} + \sigma^2_{NO\,OC})}$$

Example

- Suppose we want to design a study with 80% power to detect a mean difference of at least 5.4 mmHg between the two groups

- i.e. $\Pr\left(\left|\frac{\bar{x}_{OC} - \bar{x}_{NO\,OC}}{\sqrt{\frac{1}{n}(\sigma^2{}_{OC} + \sigma^2{}_{NO\,OC})}}\right| \geq 2\right) = 0.80$ if μ_{OC}-$\mu_{NO\,OC} \geq 5.4$mmHg

With some alegbra:

$$\Pr\left(|\bar{x}_{OC} - \bar{x}_{NO\,OC}| \geq 2 \times \sqrt{\frac{1}{n}(\sigma^2{}_{OC} + \sigma^2{}_{NO\,OC})}\right) = 0.80$$ if μ_{OC}-$\mu_{NO\,OC} \geq 5.4$mmHg

But if μ_{OC}-$\mu_{NO\,OC} \geq 5.4$mmHg, then assuming large n, $|\bar{x}_{OC} - \bar{x}_{NO\,OC}|$ is a normally distributed process with mean μ_{OC}-$\mu_{NO\,OC}$ and standard error

$$SE(\bar{x}_{OC} - \bar{x}_{NO\,OC}) = \sqrt{\frac{\sigma^2{}_{OC}}{n} + \frac{\sigma^2{}_{NO\,OC}}{n}} = \sqrt{\frac{1}{n}(\sigma^2{}_{OC} + \sigma^2{}_{NO\,OC})}$$

76

Example

- Suppose we want to design a study with 80% power to detect a mean difference of at least 5.4 mmHg between the two groups

So $\Pr\left(|\bar{x}_{OC} - \bar{x}_{NO\,OC}| \geq 2 \times \sqrt{\frac{1}{n}(\sigma^2{}_{OC} + \sigma^2{}_{NO\,OC})}\right) = 0.80$ if μ_{OC}-$\mu_{NO\,OC} \geq 5.4$mmHg

Becomes:

$$\Pr\left(\left|\frac{\bar{x}_{OC} - \bar{x}_{NO\,OC}}{\sqrt{\frac{1}{n}(\sigma^2{}_{OC} + \sigma^2{}_{NO\,OC})}}\right| \geq \frac{2 \times \sqrt{\frac{1}{n}(\sigma^2{}_{OC} + \sigma^2{}_{NO\,OC})} - (\mu_{OC} - \mu_{NO\,OC})}{\sqrt{\frac{1}{n}(\sigma^2{}_{OC} + \sigma^2{}_{NO\,OC})}}\right) = 0.8$$

77

Example

- Suppose we want to design a study with 80% power to detect a mean difference of at least 5.4 mmHg between the two groups

But on a standard normal curve, the value that cuts of 80% of the area to its right is 0.84.

So we need to solve:

$$\frac{2 \times \sqrt{\frac{1}{n}(\sigma^2{}_{OC} + \sigma^2{}_{NO\,OC})} - (\mu_{OC} - \mu_{NO\,OC})}{\sqrt{\frac{1}{n}(\sigma^2{}_{OC} + \sigma^2{}_{NO\,OC})}} = -0.84$$

Some more beautiful algebra:

$$2 \times \sqrt{\frac{1}{n}(\sigma^2{}_{OC} + \sigma^2{}_{NO\,OC})} - (\mu_{OC} - \mu_{NO\,OC}) = -0.84 \times \sqrt{\frac{1}{n}(\sigma^2{}_{OC} + \sigma^2{}_{NO\,OC})}$$

78

Example

- Suppose we want to design a study with 80% power to detect a mean difference of at least 5.4 mmHg between the two groups

Some more beautiful algebra:

$$(2+0.84)\times\sqrt{\frac{1}{n}(\sigma^2{}_{OC}+\sigma^2{}_{NO\,OC})}=(\mu_{OC}-\mu_{NO\,OC})$$

squaring both sides:

$$(2+0.84)^2\times\frac{1}{n}(\sigma^2{}_{OC}+\sigma^2{}_{NO\,OC})=(\mu_{OC}-\mu_{NO\,OC})^2$$

$$\frac{(2+0.84)^2\times(\sigma^2{}_{OC}+\sigma^2{}_{NO\,OC})}{(\mu_{OC}-\mu_{NO\,OC})^2}=n$$

79

Example

- Suppose we want to design a study with 80% power to detect a mean difference of at least 5.4 mmHg between the two groups

Plugging on our info:

$$\frac{(2+0.84)^2\times(15.3^2+18.2^2)}{(5.4)^2}=n$$

$$\frac{8.1\times 565}{29.2}=n$$

$$157\approx n$$

80

Example

- It is also possible to design a study to estimate a quantity, such as a mean difference or difference in proportions with a desired level of precision
- In other word, the necessary sample sizes can be estimated to try to get a confidence interval of a desired maximum width

81

Example

- Suppose we wanted to design a study with equal sample sizes to estimate the mean difference within ± 3 mmHg
  - i.e. design the study to have a specific precision

Now using the estimates from the small study for population SDs:

$$SE(\bar{x}_{OC} - \bar{x}_{NO\,OC}) = \sqrt{\frac{\sigma^2_{OC}}{n_{OC}} + \frac{\sigma^2_{NO\,OC}}{n_{NO\,OC}}}$$

with $n_{OC}=n_{NO\ OC}=n$ this becomes:

$$SE(\bar{x}_{OC} - \bar{x}_{NO\,OC}) = \sqrt{\frac{\sigma^2_{OC}}{n} + \frac{\sigma^2_{NO\,OC}}{n}} = \sqrt{\frac{1}{n}(\sigma^2_{OC} + \sigma^2_{NO\,OC})}$$

82

Example

- We want

$$2 \times SE(\bar{x}_{OC} - \bar{x}_{NO\,OC}) = 3$$

$$2 \times \sqrt{\frac{1}{n}(\sigma^2_{OC} + \sigma^2_{NO\,OC})} = 3$$

plugging in pilot data:

$$2 \times \sqrt{\frac{1}{n}(15.3^2 + 18.2^2)} = 3$$

Sample Data			
	n	Mean SBP	SD of SBP
OC users	8	132.8	15.3
Non-OC Users	21	127.4	18.2

83

Example

- Solving algebraically

$$2 \times \sqrt{\frac{1}{n}(15.3^2 + 18.2^2)} = 3$$

$$\sqrt{\frac{1}{n}(15.3^2 + 18.2^2)} = \frac{3}{2}$$

$$\frac{1}{n}(15.3^2 + 18.2^2) = \left(\frac{3}{2}\right)^2$$

$$\frac{(15.3^2 + 18.2^2)}{\left(\frac{3}{2}\right)^2} = n$$

$$n \approx 251$$

84

Example

- Solving algebraically

$$2 \times \sqrt{\frac{1}{n}(15.3^2 + 18.2^2)} = 3$$

$$\sqrt{\frac{1}{n}(15.3^2 + 18.2^2)} = \frac{3}{2}$$

$$\frac{1}{n}(15.3^2 + 18.2^2) = \left(\frac{3}{2}\right)^2$$

$$\frac{(15.3^2 + 18.2^2)}{\left(\frac{3}{2}\right)^2} = n$$

$$n \approx 251$$

85