Sample size determination

Brooke Anderson

9/26/2018

Study goals

We are trying to determine if the sideshow incubators can double the probability that a severely underweight infant survives.

Study design

We will study a cohort of severely underweight infants, some of whom will be cared for at the sideshow in incubators (*treated group*) and some of whom will be cared for without incubators (*control group*).

Specify a hypothesis test on a parameter (along with the underlying probability model for the data)

For each infant in the study, we will know whether it is in the treated or control group and whether it survives or dies. Let's define:

- π_1: Proportion of infants who survive in the treatment group
- π_2: Proportion of infants who survive in the control group

In this case, unit of analysis for the study is an infant. The outcome we will observe for each infant (survives $[Y = 1]$ / dies $[Y = 0]$) is expected to follow a *binomial distribution*. We now know enough to be pretty sure that our hypothesis testing will involve some kind of a test *comparing two proportions*.

Since we want to see if treatment doubles the probability of survival (i.e., has a multiplicitive, rather than additive influence on the probability of the outcome), we will conduct a hypothesis test on the *relative rate* (which we'll call λ), which has the following relationship with these proportions:

$$\lambda = \pi_1/\pi_2$$

We will conduct the following (two-sided, since we may also care about evidence that the incubators decrease the probability of survival) hypothesis test:

$$H_0 : \lambda = 1$$
$$H_1 : \lambda \neq 1$$

As a note, section 8.6 ("Testing a relative risk") in the chapter handout I sent covers sample size calculations for this type of data and hypothesis test. I am trying to use the same terminology as that section, to make it easier to follow along with how this combines with the information in the chapter. They set up the terminology for the hypothesis test above in a slightly different way, but you should be able to convince yourself that this set-up is identical to the hypothesis test they describe.

This chapter gives an equation to use for the sample size determination for this type of hypothesis test and type of data (eq. 8.14). We will not calculate this yet, but we will take a look so we'll know what other parameters we'll ultimately need:

$$n = \frac{r+1}{r(\lambda-1)^2\pi^2}\left[z_\alpha\sqrt{(r+1)p_c(1-p_c)} + z_\beta\sqrt{\lambda\pi(1-\lambda\pi)+r\pi(1-\lambda)}\right]^2$$

where:

- r is the ratio of treated to control subjects in the study (i.e., $r = 1$ if you plan on having the same number of treated and control subjects)
- λ is the effect size of scientific interest
- $\pi = \pi_2$ is the proportion of infants who survive in the control group
- z_α is a critical value that is set based on the significance level we select)
- p_c is the common proportion who survive over the two groups (see equation below to estimate this)
- z_β is a critical value that is set based on the target value of power for the test that we select

If you have equal sized groups (i.e., $r = 1$), you can estimate p_c as:

$$p_c = \frac{\pi_1 + \pi_2}{2}$$

We will use this information later to figure out all the parameters that we'll need to set to determine the required sample size, and eventually we will use this equation to calculate the sample size.

Specify the significance level α of the test

We decided on a significance level of $\alpha = 0.05$.

Now that we've set that, we can determine that:

$$z_\alpha = 1.96$$

Here's how you can deterine z_α from α in R (you're dividing the α by 2 here because we're running a two-sided hypothesis test):

```
qnorm(1 - 0.05 / 2)
```

```
## [1] 1.959964
```

You could also use a z table, which you can find online or in many statistics textbooks.

Specify an effect size that reflects an alternative of scientific interest

Let's say that we've decided that the effect size we want to test is a doubling in the probability of survival among severely underweight infants. This translates to $\lambda = 2.0$. As we discussed in the group session, you may have good reasons to have selected a different effect size of scientific interest here, which is fine as long as you justify that choice.

Obtain historical values or estimates of other needed parameters

Now, we're left with the following parameters that we still need to compute the power function of the test:

- r is the ratio of treated to control subjects in the study (i.e., $r = 1$ if you plan on having the same number of treated and control subjects)
- $\pi = \pi_2$ is the proportion of infants who survive in the control group
- p_c is the common proportion who survive over the two groups (see equation below to estimate this)

First, let's plan to have the same number of infants in the treatment and control groups ($r = 1$). This choice of $r = 1$ will result in the most power per study subject in most cases. When you decide to use a different value for r, it may be related to practical constraints on your study. For example, you many only have a limited number of incubators, which would limit the number of subjects you could treat, so in that case you may need to plan to have several infants in the control group for every infant in the treated group, which would result in a different value of r.

Next, we need to figure out a realistic value for π_2. In the group, we determined that a reasonable estimate for this is $\pi_2 = 0.2$ (i.e., 20% of infants in the control group survive).

Finally, we need to calculate p_c. This is based on both π_1 and π_2. Since we've specified that $\pi_1 = 0.2$ and that $\lambda = 2$, that requires $\pi_2 = 0.2 * 2 = 0.4$ (i.e., 40% of the treated infants survive). We can calculate p_c as (since $r = 1$):

$$p_c = \frac{\pi_1 + \pi_2}{2} = \frac{0.2 + 0.4}{2} = 0.3$$

Specify a target value of the power of the test

Let's decide to set the power to 80%. Other values would be fine, too, as long as you have some justification for your choice.

Based on this choice of power, $z_\beta = 0.8416$.

In R:

```
qnorm(0.80)
```

```
## [1] 0.8416212
```

Calculate the sample size

We now have everything we need to calculate the required sample size. In R:

```
# Define all the parameters for the equation
r <- 1
lambda <- 2.0
pi <- 0.2
z_alpha <- qnorm(1 - 0.05 / 2)
p_c <- (pi + pi * lambda) / 2
z_beta <- qnorm(0.80)

# Calculate the total required sample size (i.e., the sum of the
# required numbers in the treatment and control groups)
n <- ((r + 1) / (r * (lambda - 1) ^ 2 * pi ^ 2)) *
  (z_alpha * sqrt((r + 1) * p_c * (1 - p_c)) +
     z_beta * sqrt(lambda * pi * (1 - lambda * pi) +
                     r * pi * (1 - pi))) ^ 2
n
```

```
## [1] 162.4485
```

Therefore, we should plan for a study with at least 163 infants, half in the treatment group and half in the control group (you can round up if the total number is not even).

When we calculated this in OpenEpi, we got sample sizes of 164–182, depending on the method, so this result is consistent.

Setting up the analysis differently

The conclusions, in terms of required sample size, are approximately the same if we define the outcome as death rather than survival. In this case, $p_2 = 0.8$ (i.e., an infant in the control group has an 80% probability of dying). If an infant has twice the probability of surviving, it will have a 60% probability of dying, so in this case $\lambda = 0.6/0.8 = 0.75$ (Tori, this is where we messed up in thinking about this when we calculated this in OpenEpi).

```
# Define all the parameters for the equation
r <- 1
lambda <- 0.75
pi <- 0.8
z_alpha <- qnorm(1 - 0.05 / 2)
p_c <- (pi + pi * lambda) / 2
z_beta <- qnorm(0.80)

# Calculate the total required sample size (i.e., the sum of the
# required numbers in the treatment and control groups)
n <- ((r + 1) / (r * (lambda - 1) ^ 2 * pi ^ 2)) *
  (z_alpha * sqrt((r + 1) * p_c * (1 - p_c)) +
     z_beta * sqrt(lambda * pi * (1 - lambda * pi) +
                     r * pi * (1 - pi))) ^ 2
n
```

```
## [1] 162.4485
```

We come up with the same estimate of required sample size when we set the problem up this way.

Another example

Here's an example of using R to calculate the results from Example 8.13 from the chapter:

```
# Define all the parameters for the equation
r <- 1
lambda <- 1.4
pi <- 0.02065
z_alpha <- qnorm(1 - 0.05) # One-sided test with alpha 0.05
p_c <- (pi + pi * lambda) / 2
z_beta <- qnorm(0.90)

# Calculate the total required sample size (i.e., the sum of the
# required numbers in the treatment and control groups)
n <- ((r + 1) / (r * (lambda - 1) ^ 2 * pi ^ 2)) *
  (z_alpha * sqrt((r + 1) * p_c * (1 - p_c)) +
     z_beta * sqrt(lambda * pi * (1 - lambda * pi) +
                     r * pi * (1 - pi))) ^ 2
n
```

```
## [1] 12129.38
```