

# Course EPIB-675 - Bayesian Analysis in Medicine

## Assignment 10

In the first two questions we will consider a generic clinical trial setup, and calculate sample sizes using first standard frequentist and then Bayesian methods. In the third question, we will compare the information received from each.

The trial setup is as follows: Two drugs are being compared, A and B. The main outcome is dichotomous, which we will generically call “success” and “failure”. The prior information is such that Drug A is 95% certain to have a success rate between 60% to 80%, while Drug B is 95% certain to have a success rate between 70% and 85%. It is desired to estimate the true difference between the success rates of A and B to an accuracy of  $\pm 2.5\%$  (total interval width 5%) with a 95% interval.

1. Use standard frequentist methods to calculate the sample size required for each group. Assume equal sized groups, and use the midpoints of the prior intervals given above as point estimates. You can plug your numbers into the sample size formula below

$$N_{per\ group} = \frac{z_{1-\alpha/2}^2(p_1(1-p_1) + p_2(1-p_2))}{d^2} = \frac{4z_{1-\alpha/2}^2(p_1(1-p_1) + p_2(1-p_2))}{w^2}$$

where  $p_1$  and  $p_2$  are the estimates of the success rates in groups 1 and 2, respectively,  $z_{1-\alpha/2}$  is the usual normal percentile (1.96 for a 95% interval), and  $w = 2d$  is the total length of the desired interval. You can check your answer using the frequentist sample size calculator found here (very easy to download and install, and trivially easy to use):

<http://www.medicine.mcgill.ca/epidemiology/Joseph/software/SampleSizeCalculator/samplesize.zip>

2. We will now redo the sample size calculations using a variety of Bayesian sample size criteria.

(a) The first step is to convert the prior information on  $p_1$  and  $p_2$  given above into beta prior distributions. Use the mid-points of the prior intervals (the same values you used in you used in question 1) as the means of beta distributions, and divide the ranges of the prior intervals by 4 to get approximate SD's. Using the formulae given below (we have seen them previously, in the section of the course on inference in binomial proportions)

$$\alpha = -\frac{\mu (\sigma^2 + \mu^2 - \mu)}{\sigma^2}$$

and

$$\beta = \frac{(\mu - 1) (\sigma^2 + \mu^2 - \mu)}{\sigma^2}$$

find the appropriate beta density for each range.

(b) Install and load in the R package called SampleSizeProportions, and use it to fill in the following table of sample sizes (again, assume equal sized groups, first row is done for you as an example, see code below):

Bayesian sample size Criterion	Sample Size
Average Coverage Criterion (ACC)	2242
Average Length Criterion (ALC)	
Worst Outcome Criterion (WOC)	
MBL Average Coverage Criterion (MBLACC)	
MBL Average Length Criterion (MBLALC)	
MBL Worst Outcome Criterion (MBLWOC)	
Modified Worst Outcome Criterion (MODWOC(prob = 0.5))	
Modified Worst Outcome Criterion (MODWOC(prob = 0.9))	
MBL Modified Worst Outcome Criterion (MBLMODWOC(prob = 0.5))	
MBL Modified Worst Outcome Criterion (MBLMODWOC(prob = 0.9))	

Hints:

- To fill in the first line of the table, you need to type something like

```
> propdiff.acc(.05, 58.1, 24.9, 95.325, 27.675)
[1] 2242 2242
```

so that 2242 subjects are needed in each group here (your result may vary a bit, as a Monte Carlo algorithm is used within the program, but

should be very close to 2242). Note that the program may take a few seconds or even a few minutes to run, depending on the speed of your computer (takes about 1 second on a 3.2 GHz machine).

2. Refer to the article on the course web page on Bayesian sample size for details about how to interpret each criterion.

3. We will now compare the information available from each of these sample size calculations. How do the sample sizes from the frequentist approach calculated in question 1 compare to the Bayesian approaches calculated in question 2? Take some examples of Bayesian sample sizes that are larger/smaller than the frequentist approach, and explain why they are larger or smaller.

In order to answer the next two questions, you will need to download the Diagnostic test program for one and two tests from the course web page. You can just cut and paste them directly into R, or save and use the source command to read them in.

Look at the example file called `tt2_example.txt` for hints about how to run the programs and interpret the output.

4. In this problem we will analyse diagnostic test data for a single diagnostic test. Suppose we are looking to estimate the prevalence of a certain disease, and thus give a diagnostic test for that condition to 500 subjects. Suppose that 150 of these subjects test positive, and the other 350 test negative.

(a) Assuming the test to provide a perfect indication of the presence of absence of the disease (i.e., assume the test is a perfect gold standard), what would be your estimate of disease prevalence? Provide a CI around this point estimate (use standard frequentist binomial methods for this question. For example, you can use the R function `prop.test`).

(b) Now assume that the sensitivity is known to be exactly 70%, and the specificity is known to be exactly 95%. We know that in this case the probability of a positive test is given by

$$p = \pi * S + (1 - \pi) * (1 - C)$$

where  $p$  is the probability of a positive test,  $\pi$  is the prevalence, and  $S$  and  $C$

are the sensitivity and specificity, respectively. To get a point estimate and confidence interval for  $\pi$ , you need to solve the above equation for  $\pi$  (a few algebra steps), plug in the known values for  $S$  and  $C$ , and then plug in the point estimate, and lower and upper limits for  $p$  to get the point estimate, and lower and upper limits for  $\pi$ .

(c) Finally (and most realistically), assume that the sensitivity and specificity are not known exactly, but only up to an interval. In particular, assume that nothing is known a priori for the prevalence (i.e., assume a  $\text{beta}(1,1)$  prior for the prevalence), but assume prior distributions of  $\text{beta}(58.1,24.9)$  for the sensitivity (which has mean of 70% as in part (b), but with 95% interval approximately [60%, 80%]) and  $\text{beta}(71.25, 3.75)$  for the specificity (which has mean of 95% as in part (b), but with 95% interval approximately [90%, 100%]). Instructions for running the program required is given at the top of the function itself, which you can look at by typing its name in R.

Basically, you need to run a command that looks like:

```
tt1.gibbs(tp, tn, astart, cstart, sensstart, specstart, prevstart,  
  alphaprev, betaprev, alphasens, betasens, alphaspec, betaspec,  
  size, throwaway, skip)
```

where each parameter input is described in the function. Use any starting values you wish that are in the correct range for each variable. Run at least 20,000 iterations, which should not take too long.

(d) Compare the three point and interval estimates for the prevalences calculated in parts (a), (b), and (c). Does accounting for imperfect sensitivity and specificity have much effect on prevalence estimates (comparing estimates from (a) to those in (b))? Does accounting for imperfect knowledge about the sensitivities and specificities have much effect on prevalence estimates (comparing estimates from (b) to those in (c))? Note that parts (a) and (b) can be done from either frequentist or Bayesian viewpoints, but part (c) can only be done by Bayesian methods.

5. To demonstrate that it can be just as easily done, in this problem we will estimate the prevalence when two imperfect diagnostic tests are applied to a sample. Here we will reproduce the estimates in the *Strongyloides* example given in the article on the course web page. Using the prior distributions

for the sensitivity and specificity of the stool examination and the sensitivity and specificity of serological testing, and using a  $\text{beta}(1,1)$  prior distribution for the prevalence, and using the two-by-two set of data given in the article, report the posterior distribution for the prevalence. Provide the lines of programming that you need to type in R in order to run this analysis, and the posterior median and 95% interval for the prevalence.