Design and Analysis of Clinical Trials

Assignment #6

Due 11/29/2022

1. **A randomized clinical trial will compare surgical procedures A and B for the repair of a torn ankle ligament. The primary outcome was recovery of the ankle ligament post-surgery (scored as Excellent vs. Not Excellent). The randomization assignments were stratified by whether the patient had previous ligament damage. Results are below:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Outcome** | **Prior Damage** | | **No Prior Damage** | |
| **Procedure A** | **Procedure B** | **Procedure A** | **Procedure B** |
| Excellent | 6 | 3 | 17 | 9 |
| Not Excellent | 14 | 17 | 13 | 21 |

**Using methods appropriate to this setting:**

* 1. **Provide a point estimate and 95% CI for the odds ratio of an excellent outcome under procedures A vs. B.**

I will construct a 2x2 table from the table above that reflect treatment A vs B. I will use Fisher’s Exact Test in R to answer this question:

dat = c(23,27,12,38)

mat\_1a = matrix(dat,2,2,)

rownames(mat\_1a) = c("Excellent", "Not Excellent")

colnames(mat\_1a) = c("A", "B")

fisher.test(mat\_1a)

# Fisher's Exact Test returns an odds ratio of 2.6975, which matches the

# odds ratio when calculated => (23/27) / (12/38) = 2.6975

#

# We also find that the 95% confidence interval is [1.0621, 6.9949], also

# from the fisher.test() function.

* 1. **Test whether the rates of excellent outcomes are different between A and B at a 5% significance level.**

H0=proportion of excellent outcomes are the same between procedures A & B.

HA=proportion of excellent outcomes are different between procedures A & B.

For this, I used R to find the p\_value to this hypothesis using Fisher’s test and the Chi-Squared Test:

fisher.test(mat\_1a)

chisq.test(mat\_1a)

# We find a p\_value of 0.036 for this hypothesis. Thus, we reject H\_0 and conclude that there is a

# difference of proportion of outcomes between A and B at a significance of 5%.

1. **A clinical trial is being developed to compare efficacy of a new cholesterol reducing drug to placebo. The primary outcome is reduction in cholesterol level from randomization to 4 weeks post-randomization. The investigator wants to have 80% power to detect an average cholesterol reduction of 5 units for the new drug compared to 0 units for placebo, using a two-sided 5% significance level. The SD of the reduction is expected to be 10. A group sequential design with three equally spaced analyses is planned.**

Observe the following R code:

Library(gsDesign)

power=0.8

alpha=0.05

delta=5

strd\_dev = 10

var = strd\_dev^2

z\_halpha = abs(qnorm(.025)) #=1.96

z\_beta = qnorm(0.8) #=.84

group = 2

looks = 3

* 1. **Find a Pocock stopping boundary that has the desired operating characteristics. Describe the decision rule for all stages, the maximum sample size targeted, and the maximum information needed.**

From my R code:

# Problem 2 (A)

# Using table 14.2 of Piantadosi

pocock = 2.289

# If the Z\_score is ever greater than this value, we stop the trial. This is the Pocock stopping boundary for 3 analyses.

look = c(1,2,3)

Pocock\_tab = c(2.289,2.289,2.289)

DecisionRule = data.frame(look,Pocock\_tab)

OUTPUT:

look Pocock\_tab

1 1 2.289

2 2 2.289

3 3 2.289

# In order to find the maximum sample size, we must first find the inflation factor, R\_p.

# We can find the inflation factor, R, from table 2.2 and 2.4 of Jennison and Turnbull:

R\_p = 1.166 #R\_p(3, 0.05, 0.8)

n\_f = (2\*(z\_halpha + z\_beta)^2\*var) / delta^2

n\_max = R\_p \* n\_f #PER GROUP

# Thus, n\_max = 73.2 per arm.

# Group size to enroll for each stage is m=73.2 / 3 = 24.4 ~=~ 24 per treatment

# Stop and reject H\_0 at analysis j if |Z(t\_j)| >= 2.289 for j=1,2,3.

# Then, the maximum information:

n1 = n\_max

n2 = n1

MI\_p = (var\*(1/n1 + 1/n2))^(-1)

# We get MI = 0.366

* 1. **Perform the same steps in part (A) using an O’Brien-Fleming stopping boundary.**

## We have 3 equally spaced analyses planned.

## According to the gsDesign function -> gsDesign(k=3, test.type=2, alpha=.025,beta=.2, sfu=”OF”,n.fix=n\_f)

OBF1 = 3.47

OBF2 = 2.45

OBF3 = 2.00

## Then, we have the decision rule for all stages below:

OBF\_tab = c(OBF1, OBF2, OBF3)

RuleOBF = data.frame(look, OBF\_tab)

OUTPUT:

look OBF\_tab

1 1 3.47

2 2 2.45

3 3 2.00

# If any |Z(t\_j)| are greater than these values at the time of their look, we stop and reject H\_0 at analysis.

## Maximum sample size targeted

R\_b = 1.024 #R\_b(3, 0.05, 0.8)

n\_maxOBF = R\_b \* n\_f #This is per group

## n\_maxOBF = 64.298 per group.

## Group size to enroll for each stage is m=64.298/3 = 21.099 ~=~ 21 per treatment

## Stop and reject H\_0 at analysis j if |Z(t\_j)| >= 1.985\*(sqrt(3/j)) for j=1,2,3.

## With maximum information for 2 groups:

n1b = n\_maxOBF

n2b = n1b

MI\_b = (var\*(1/n1b + 1/n2b))^(-1)

# Maximum Information for OBF is 0.3214

* 1. **Which design do you prefer: (A) or (B)? Why?**

Looking at both designs, I would prefer the O'Brien-Fleming design for a few reasons:

Firstly, having a harder stopping boundary early would guard against early un-representative behavior in the data. Early outlier data could cause a stop in the Pocock design. This also piggybacks into a more sensitive stopping boundary later, since more data should be more reflective of the true efficacy and should be more resistant to a narrower stopping boundary.

Secondly, OBF offers a smaller expected sample size. This makes interim analysis easier in practice.

* 1. **Suppose the t-statistics at the three interim analyses were 1.37, 2.12, and 2.17. For each of the designs (A) and (B), when would early stopping occur (if at all) and what would be the final conclusion of the study?**

t\_stat = c(1.37, 2.12, 2.17)

prob2d = data.frame(look, Pocock\_tab, OBF\_tab, t\_stat)

plot(look, Pocock\_tab, type='l')

lines(OBF\_tab)

lines(t\_stat)

For these t-statistics, we would find that for design (A), the Pocock design, we would get no stoppage because of the t-statistics being smaller than the pocock limit = 2.289 in all looks. The conclusion of this interim analysis would be to NOT stop the trial early.

For design (B), the OBF design, we would get stoppage at look 3 as a result of the third, and most sensitive, boundary being greater than the final of the t-statistics. The conclusion of the interim analysis for design (B) would be to stop the trial early.

1. **Consider a trial comparing response rates from two treatment arms, i.e. testing . Design a group sequential study with 90% power to detect an improvement in the success rate from 20% in the control arm to 35% in the treatment arm. Assume a 5% two-sided significance level and an O’Brien-Fleming boundary with 4 total analyses. Describe the stopping rule, the maximum sample size, and the maximum information needed.**



# 2 treatment arms

# H\_0: delta = pi\_1 - pi\_2 = 0.

pwr = .1

alpha = 0.025 #two-sided significance for alpha=.05

delta=.35-.20 #20% success in control to 35% success in treatment

looks = 4 #O'Brien-Fleming boundary

groups = 2 #control arm and treatment arm

varp3 = 1 #Since this is a binomial proportion

z\_a = abs(qnorm(.025))

z\_b = abs(qnorm(.9))

#From the Jennison & Turnbull Table

R\_bp3 = 1.022

n\_fp3 = (2\*(z\_a+z\_b)\*varp3)/delta^2

gsDesign(k=4, test.type=2,alpha=alpha,beta=pwr,sfu="OF",n.fix=n\_fp3)

OUTPUT:

Symmetric two-sided group sequential design with

90 % power and 2.5 % Type I Error.

Spending computations assume trial stops

if a bound is crossed.

Analysis N Z Nominal p Spend

1 74 4.05 0.0000 0.0000

2 148 2.86 0.0021 0.0021

3 221 2.34 0.0097 0.0083

4 295 2.02 0.0215 0.0145

Total 0.0250

n\_maxp3 = R\_bp3 \* n\_fp3 # = 294.47 ~=~ 295 total and 74 per treatment

## Group size to enroll for each stage is m=295. 295/4 = 74 per treatment.

## Stop and reject H\_0 at analysis j if |Z(t\_j)| >= 2.02\*(sqrt(4/j)) for j=1,2,3,4.

n1p3=n\_maxp3/2

n2p3=n1

MI\_p3 = (varp3\*(1/n1+1/n2))^(-1)

#Gives a maximum information of 36.607

**\* Please submit any code used to perform the calculations for your solutions.**