Analyzing RCTs: A Cookbook

Inge Christoffer Olsen

2019-08-09

Contents

1 Preface **5**
1.1 Structure of the book 5
1.2 Acknowledgements 5

2 Introduction **7**
2.1 Trial Flow 7
2.2 Simulated dataset 7

3 Continuous endpoints **9**
3.1 Single follow-up 9
3.2 Repeated follow-up 9

4 Dichotomous endpoints **11**
4.1 Single follow-up 11

5 Applications **15**
5.1 Example one 15
5.2 Example two 15

6 Final Words **17**

Chapter 1

Preface

Placeholder

1.1 Structure of the book

1.2 Acknowledgements

Chapter 2

Introduction

Placeholder

2.1 Trial Flow

2.2 Simulated dataset

Chapter 3

Continuous endpoints

Placeholder

3.1 Single follow-up

3.1.1 Stata code

3.1.2 R code

3.1.3 Reporting

3.2 Repeated follow-up

3.2.1 Simple model

3.2.2 Model with treatment-time interaction

3.2.3 Model with treatment-time interaction and baseline information

Chapter 4

Dichotomous endpoints

4.1 Single follow-up

For a single follow-up assessment of a dichotomous endpoint, the main method I use is a standard logistic regression. Then we can adjust for stratification factors in the randomisation in addition to other pre-specified covariates, both categorical and continuous. In the simulated example, we define that the primary outcome is the dichotomous categorical outcome at time 3. Note that usually the baseline status of all patients are negative for the outcome, so adjusting for baseline is not necessary.

4.1.1 Stata code

```
use "stata/rct", clear
tabulate catout trt
logistic catout i.trt i.site covar  if time==3, coef
```

```
##
## . use "stata/rct", cle(all strata combined)
##
## . tabulate catout trt
##
## Categorica |       Treatment
##  l outcome |   Placebo     Active |     Total
## -----------+----------------------+----------
##   Negative |        95        131 |       226
##   Positive |       105         61 |       166
## -----------+----------------------+----------
##      Total |       200        192 |       392
```

```
##
## . logistic catout i.trt i.site covar  if time==3, coef
##
## Logistic regression                               Number of obs     =         98
##                                                   LR chi2(5)        =      48.59
##                                                   Prob > chi2       =     0.0000
## Log likelihood = -36.862204                       Pseudo R2         =     0.3973
##
## ------------------------------------------------------------------------------
##       catout |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
## -------------+----------------------------------------------------------------
##          trt |
##       Active |  -2.890301   .7850252    -3.68   0.000    -4.428922   -1.351679
##              |
##         site |
##            2 |   .7783404   .8580245     0.91   0.364    -.9033566    2.460037
##            3 |   1.423791   .7786531     1.83   0.067    -.1023412    2.949923
##            4 |   .0253234   .8082887     0.03   0.975    -1.558893     1.60954
##              |
##        covar |   1.001078   .2329461     4.30   0.000     .5445124    1.457644
##        _cons |  -2.463577   .8925892    -2.76   0.006     -4.21302   -.7141344
## ------------------------------------------------------------------------------
##
## .
```

Note that the use the `coef` option to get the log odds ratio estimates.

4.1.2 R code

```
rct <- read_dta("stata/rct.dta") %>%
  modify_at(c("trt","catout"), haven::as_factor, levels = "labels") %>%
  modify_at(c("site","time"), haven::as_factor)
rct %>%
  filter(time==3) %>%
  glm(catout ~ trt + site + covar , data=., family = binomial) %>%
  summary
```

```
##
## Call:
## glm(formula = catout ~ trt + site + covar, family = binomial,
##     data = .)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.9659  -0.5335   0.1943   0.5095   2.8873
```

```
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.46358    0.89259  -2.760 0.005780 **
## trtActive   -2.89030    0.78502  -3.682 0.000232 ***
## site2        0.77834    0.85802   0.907 0.364337
## site3        1.42379    0.77865   1.829 0.067470 .
## site4        0.02532    0.80829   0.031 0.975007
## covar        1.00108    0.23295   4.297 1.73e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 122.318  on 97  degrees of freedom
## Residual deviance:  73.724  on 92  degrees of freedom
## AIC: 85.724
##
## Number of Fisher Scoring iterations: 6
```

Chapter 5

Applications

Some *significant* applications are demonstrated in this chapter.

5.1 Example one

5.2 Example two

Chapter 6

Final Words

We have finished a nice book.