CPH675-FINAL

Dominic LaRoche

Tuesday, May 05, 2015

# Question 1

* 1. [8 pts] You are part of a research team that is going to carry out a 2x2 factorial study investigating Tai Chi and meditation to reduce anxiety in cancer survivors. Create a randomization list for this trial which is based on simple randomization, but is equally balanced across the arms. Start from first principles by using a statistical distribution. The total sample size is 400. Show your code and appropriate table(s) that demonstrate that you have achieved balance.

trtA<-trtB<-trtC<-trtD<-vector()  
trtGroup<-character()  
i<-0  
while(length(trtGroup)<400 ){  
 u<-runif(1)  
 if(u < (1/4) & length(trtA)<100){  
 i<-i+1  
 trtGroup[i]<-"Treatment A"  
 trtA<-c(trtA,1)  
 }  
 if(u > (1/4) & u < (2/4) & length(trtB)<100){  
 i<-i+1  
 trtGroup[i]<-"Treatment B"  
 trtB<-c(trtB,1)  
 }  
 if(u > (2/4) & u < 3/4 & length(trtC)<100){  
 i<-i+1  
 trtGroup[i]<-"Treatment C"  
 trtC<-c(trtC,1)  
 }  
 if(u > (3/4) & length(trtD)<100){  
 i<-i+1  
 trtGroup[i]<-"Treatment D"  
 trtD<-c(trtD,1)  
 }  
}  
summary(factor(trtGroup))

## Treatment A Treatment B Treatment C Treatment D   
## 100 100 100 100

* 1. [2 pts] The above randomization is neither stratified nor blocked. Discuss the implications of this.
* Since the above randomization is not stratified it is more likely that the final allocation will have imbalance on potentially important covariates. Since there is no blocking this imbalance could be especially pronounced early in the trial when the sample size is small. Without blocking it is also possible that the allocation between the arms could become imbalanced in the middle of the trial.
  1. [4 pts] Discuss some of the issues (both positive and negative) that arise due to this design and intervention structure (do not discuss the likelihood of the intervention being effective or not).
* The factorial design can be more efficient if it is assumed there is no interaction between Tai Chi and meditation. However, if the study is powered based on this assumption (to reduce sample size) then the power to detect an interaction will be low, particularly if the interaction is small. If an interaction is present then the power for detecting main effects is reduced because cells can no longer be collapsed. Conversely, if an interaction is expected the study can be powered specifically to test the interaction.
  1. [8pts] Anxiety is measured as a binary variable, and is measured once only (post-intervention). The primary objective is to compare both the active interventions to control. A secondary hypothesis is that men and women may respond differently to the intervention. Write a brief statistical analysis plan for this study.
* We will use logistic regression to test for an interaction between Tai Chi and meditation. If we find no evidence for an interaction we will collapse treatment combination groups and compare those that received the Tai Chi intervention to those that did not (control only group and meditation + control group) and those with meditation to those that did not receive the meditation intervention using tests. We will use a Mantel-Haeztel test for homogeneity of odds ratios to test for a difference in treatment effect between men and women. If we find evidence of an interaction we will perform the same tests as outlined above but without collapsing groups.

# Question 2

* 1. [8 pts] Find the sample size required for a three armed trial where allocation ratio is 2:1 for each of the two active interventions compared to the control (2:2:1). Assume 80% power, type I error rate of 0.05, two-sided tests. The researchers aim to find a standardized difference between the two active arms of 0.3, and a standardized difference of at least 0.5 between the control group and either of the active arms.
* We will need to find the minimum sample size such that we have 80% power for both the comparison between the two arms or between each arm and the control:

ssize<-function(alpha, beta, sigma, k, margin){  
 n2<- (qnorm(1 - alpha/2) + qnorm(1 - beta))^2 \* sigma^2 \* (1 + 1/k)/margin^2  
 n1<-k\*n2  
 return(list(N=ceiling(n1)+ceiling(n2),N1=ceiling(n1),N2=ceiling(n2)))  
}  
#SS for between the two treatments  
t1Vt2<-ssize(alpha=.05,beta=.2,sigma=1,k=1,margin=.3)  
t1Vt2

## $N  
## [1] 350  
##   
## $N1  
## [1] 175  
##   
## $N2  
## [1] 175

#ss for between either treatment arm and the control  
tVc<-ssize(alpha=.05, beta=.2,sigma=1, k=2, margin=.5)  
tVc

## $N  
## [1] 143  
##   
## $N1  
## [1] 95  
##   
## $N2  
## [1] 48

The minimum required sample size for finding a difference between the two treatment arms is 175 samples per arm. The minimum required sample size for finding a difference between either treatment and the control is 95 samples per treatment arm. Therefore, if we power for the between arm comparison we will have adequate power for the comparison with control. Based on the fixed allocation ratio of 2:2:1 a sample size of 175, 175, and 87 will have 80% power at the for either comparison. This sample size could be reduced if the allocation ratio could be changed.

* 1. [4 pts] A secondary outcome is binary. Do they have enough power (= 80%) to detect a relative risk of 1.5 or greater between the control group and either of the active arms? State all the assumptions you are making.

rrsize<-function(alpha=.05, beta=.2, k, pt, pc){  
 or=(pt\*(1-pc))/(pc\*(1-pt))  
 n.c <- (qnorm(1 - alpha/2) + qnorm(1 - beta))^2 \* (1/(k\*pt\*(1 - pt)) + 1/(pc\*(1 - pc)))/(log(or))^2  
 n.c  
}  
#determine minimum proportion in controls that would give 80% power to detect a RR of 1.5  
x<-seq(.2,.4,.001)  
ss<-rrsize(k=2,pt=1.5\*x, pc=x)  
min(x[which(ss<87)])

## [1] 0.371

* There is insufficient information to make this determination. The sample size (and the power function) for the equivalence of two proportions is dependent on the absolute proportions given:

where and represent the proportions in the control and treatment groups respectively (from the text). With a relative risk of 1.5 we can substitute for but the sample size will still depend on , which is unknown in this case. It can be shown (taking the limit as approaches 0 using L'Hopital's rule) that the sample size must increase as decreases\*. The smallest proportion in controls that would allow for 80% power to detect a relative risk of 1.5 with the given sample size (treatment = 175 and control = 87) is approximately 0.37. Under this scenario the proportion under treatment would be . \*(The same applies as approaches 1 and the upper limit of can be found in a similar way, but I assume a )

# Question 3

### Download the data RCT\_wide from D2L to answer the following questions.

* 1. [3 pts] This is a longitudinal study, which means that the data are not independent. What are three methods which could account for the lack of independence of the data?
     1. Collapse data for each individual to get a single change score and model only these.
     2. Use a generalized estimating equation (GEE)
     3. Use a mixed effects model with a random intercept for each individual
  2. [3 pts] Create an informative table of missing data

Table 1. Proportion of missing responses by treatment arm and time period.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Time 1 | Time 2 | Time 3 | Time 4 | Time 5 |
| **Control** | 0 | 0.09 | 0.16 | 0.26 | 0.33 |
| **Treatment** | 0 | 0.06 | 0.09 | 0.14 | 0.2 |

* 1. [2 pts] You should note that the missing data rates differ by group. Does this give evidence that the data are MCAR, MAR, or MNAR? Explain briefly.
* Since the data appear to differ by group this gives some evidence that the data are Missing at Random (MAR) or Missing Not at Random (MNAR). Conditioning on treatment group the data may be missing completely at random, i.e. it may be unrelated to the unknown outcome but rather the treatment regimen. However, it may also be that the control group had worse outcomes and this is the reason why the data are missing. In this case the missingness is related to the unobserved outcome and is therefore MNAR.

### For problems d) and e), use a means mixed model, assuming a compound symmetric variance-covariance matrix.

* 1. [2 pts] Is there a difference in the pattern of change over time by arm? Explain.

rct\_long<-reshape(rct,varying=paste0("y",1:5), v.names="y",times=1:5,idvar=c("id","group"), direction="long")  
library(lme4)  
library(lmerTest)  
m1<-lmer(y~factor(group)\*factor(time) + (1|id), data=rct\_long)  
lmerTest::summary(m1,ddf="Kenward-Roger")$coef

## Estimate Std. Error df t value  
## (Intercept) 77.0397877 1.333832 343.0582 57.7582239  
## factor(group)1 0.7341537 1.886324 343.0582 0.3891981  
## factor(time)2 -2.5965700 1.152237 675.6655 -2.2532430  
## factor(time)3 -4.6413420 1.185784 679.7170 -3.9135563  
## factor(time)4 -4.6271161 1.238058 682.1666 -3.7367244  
## factor(time)5 -7.3709998 1.280062 682.7753 -5.7572249  
## factor(group)1:factor(time)2 0.2388602 1.618544 673.1055 0.1475631  
## factor(group)1:factor(time)3 2.1321351 1.651712 676.0999 1.2907035  
## factor(group)1:factor(time)4 2.0151966 1.704459 678.1303 1.1821369  
## factor(group)1:factor(time)5 4.5431477 1.753934 678.8809 2.5898599  
## Pr(>|t|)  
## (Intercept) 8.252514e-179  
## factor(group)1 6.973711e-01  
## factor(time)2 2.456370e-02  
## factor(time)3 1.001009e-04  
## factor(time)4 2.020400e-04  
## factor(time)5 1.292466e-08  
## factor(group)1:factor(time)2 8.827318e-01  
## factor(group)1:factor(time)3 1.972477e-01  
## factor(group)1:factor(time)4 2.375657e-01  
## factor(group)1:factor(time)5 9.807414e-03

* There is a difference in the pattern of change by arm. The treatment group has a higher mean by 4.54 points (95% CI 1.11, 7.97) at the last time point.
  1. [6 pts] Use a contrast within your mixed model to find the difference between groups in the change from baseline at each post-baseline time-point, and the associated 95% confidence intervals. Use a table to display your results, and show your code.

Table 2. The difference between the treatment and control arms in the change from baseline at each time-point (difference = treatment-control).

library(contrast)  
library(multcomp)

## Loading required package: mvtnorm  
## Loading required package: TH.data

rct\_long$fgrp<-factor(rct\_long$group)  
rct\_long$ftm<-factor(rct\_long$time)  
rct\_long$fsex<-factor(rct\_long$sex)  
m1b<-lm(y~fgrp\*ftm,data=rct\_long)  
  
cct1<-contrast::contrast(m1b,a=list(fgrp="1",ftm="1"),b=list(fgrp="0", ftm="1"))  
time1<-unlist(summary(glht(m1,linfct=cct1$X))$test[3:4])

## Loading required package: lmerTest  
## Loading required package: Matrix  
## Loading required package: lme4  
## Loading required package: Rcpp  
##   
## Attaching package: 'lmerTest'  
##   
## The following object is masked from 'package:lme4':  
##   
## lmer  
##   
## The following object is masked from 'package:stats':  
##   
## step

cct1$X[1:10]<-c(0,0,0,0,0,0,1,0,0,0)  
time2<-unlist(summary(glht(m1,linfct=cct1$X))$test[3:4])  
  
cct1$X[1:10]<-c(0,0,0,0,0,0,0,1,0,0)  
time3<-unlist(summary(glht(m1,linfct=cct1$X))$test[3:4])  
  
cct1$X[1:10]<-c(0,0,0,0,0,0,0,0,1,0)  
time4<-unlist(summary(glht(m1,linfct=cct1$X))$test[3:4])  
  
cct1$X[1:10]<-c(0,0,0,0,0,0,0,0,0,1)  
time5<-unlist(summary(glht(m1,linfct=cct1$X))$test[3:4])  
  
out2<-as.data.frame(rbind( time2, time3, time4, time5))  
out2$LCI<-out2$coefficients.1-1.96\*out2$sigma.1  
out2$UCI<-out2$coefficients.1+1.96\*out2$sigma.1  
rownames(out2)<-paste("Time",2:5, sep=" ")  
names(out2)<-c("Estimated difference", "SE", "Lower 95% CI", "Upper 95% CI")  
pandoc.table(out2,style="rmarkdown", split.table=Inf, round=2)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimated difference | SE | Lower 95% CI | Upper 95% CI |
| **Time 2** | 0.24 | 1.62 | -2.93 | 3.41 |
| **Time 3** | 2.13 | 1.65 | -1.11 | 5.37 |
| **Time 4** | 2.02 | 1.7 | -1.33 | 5.36 |
| **Time 5** | 4.54 | 1.75 | 1.11 | 7.98 |

I used the model:

The mean scores for the control group and treatment group at time 2 is therefore:

Subtracting the mean score for each group at baseline to get the mean change in score we get:

To get the difference we subtract the mean change in score at time 2 for the control group from the treatment group:

This can repeated for each time point to get the remaining estimates.

* 1. [6 pts] Suppose that the primary outcome was the comparison of groups in their change from baseline at the final time-point. Write your results in a few sentences that would be appropriate for a journal.
* We used a linear mixed effects model with random intercepts for each individual to account for the repeated measures. We found a statistically significant difference in the change in mean score between the two arms. The treatment group mean dropped 4.54 points less (95% CI 1.11, 7.98) than the control group mean leading to a higher overall score at the final time-point. All calculations were performed using R version 3.1.1 and the lme4, contrast, and glht packages.
  1. [2 pts] A secondary objective was to investigate whether the difference between groups in the patterns of change over time was modified by sex. Test this hypothesis.

#use a nested model comparison to determine if sex modifies the time\*trt interaction  
m2<-lmer(y~fgrp\*ftm+fsex\*fgrp + fsex\*ftm + (1|id), data=rct\_long)  
m3<-lmer(y~fgrp\*ftm\*fsex + (1|id), data=rct\_long)  
#improvement in model fit from just the group\*time\*sex interaction  
anova(m2,m3)

## refitting model(s) with ML (instead of REML)

## Data: rct\_long  
## Models:  
## object: y ~ fgrp \* ftm + fsex \* fgrp + fsex \* ftm + (1 | id)  
## ..1: y ~ fgrp \* ftm \* fsex + (1 | id)  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)  
## object 18 6350.5 6436.3 -3157.3 6314.5   
## ..1 22 6354.2 6459.0 -3155.1 6310.2 4.299 4 0.3671

* I did not find a statistically significant modification of the patterns of change over time by sex (p= 0.4)
  1. [2 pts] How could you improve the precision of your results?
* Including relevant covariates could improve model fit and increase the precision of the estimates. For example, including an interaction between sex and time reduced the standard error of the time by treatment interaction estimate.

# Question 4: Short answer, 2 pts each

* 1. What method should be used to test the difference in overall survival between two arms in an RCT?
* A cox proportional hazards model or a log-rank test.
  1. A two-arm parallel RCT found a hazard ratio of 1.14 ( 95%CI, 1.10-1.17) for the outcome of survival when comparing a new intervention to usual care. Explain what this means.
* This tells us that the participants on the treatment arm had significantly lower survival times than those on usual care.
  1. State two of the functions of a data safety monitoring board.
     1. Monitor the number of adverse events in both arms during the trial
     2. Monitor the protocol adherence and alert the PI of data quality is low
  2. Suppose a trial uses a block size of 8 during randomization. What is the largest sample size imbalance between the arms that could occur during the trial?
* With a block size of eight the largest imbalance possible is 4.
  1. During an interim analysis for a trial that was half way through recruitment, it was found that sex, which was thought to be highly prognostic of outcome, was imbalanced between the arms. What are two approaches that could be used to handle this imbalance?
     1. Do nothing. The randomization process should work for a large enough study (although at a minimum it would be wise to check for a systematic problem with the randomization schema)
     2. Enrich the trial for the sex that is under-represented.
  2. Suppose a cluster randomized trial does not take between cluster variation into account. Will this affect bias, type I error rate, both or neither?
* This will affect the type 1 error rate since the variance of the estimates will be underestimated.
  1. The following is output from an individual level analysis of a cluster randomized trial. There were 50 hospitals randomized, with 20 patients at each hospital.
  2. What was the empirical (observed) ICC?
  3. What was the empirical design effect?
  4. Suppose you are planning a group sequential trial with 2 intermediate looks. Using the O’Brien-Fleming alpha spending function, find the alpha required to stop the trial early at both stages.
* Adding two intermediate looks will give a total of 3 sequential groups and 3 tests. Assuming equal group size:

OF<-function(t, alpha=.05){  
 Alpha<-2-2\*pnorm(abs(qnorm(alpha/2)/sqrt(t)))  
 return(Alpha)  
}  
first.look<-OF(1/3)  
second.look<-OF(2/3)-first.look

To stop the trial early at the first look would require an . To stop the trial early at the second look would require .

* 1. Briefly explain the concept of stopping for futility
* If there is only a very small probability that the trial will be able to show efficacy if it were to continue to completion then the trial can be stopped early for futility to avoid exposing participants to unnecessary research.
  1. suppose you are carrying out a non-inferiority trial, where lower scores are better and the non-inferiority margin is 4. The confidence interval is (-3.2, 1.9). Do you reject or fail to reject the null hypothesis?
* We reject the null hypothesis of inferiority because the upper confidence limit on the difference between the new and old treatment is below the confidence limit of 4.