Conditional Inference for Repeated Measures models

Livio Finos

Contents

1 Introduction

knitr::opts_chunk\$set(echo = TRUE)

1.1 The data

(Fictitious data)

ERP experiment

- 20 Subjects,
- 6 Channels: O1, O2, PO7, PO8, P7, P8
- Stimuli: pictures. Conditions:
	- **–** 1 (f): fear (face)
	- **–** 2 (h): happiness (face)
	- **–** 3 (d): disgust (face)
	- **–** 4 (n): neutral (face)
	- **–** 5 (o): object
- Measure: Area around the component P170

Setting parameters, importing the data:

```
rm(list=ls())
library(flip)
#
```

```
# # example of files contents:
# # s01 NC P7 f -7.1121
# # s01 NC P7 h -7.2582
# # s01 NC P7 d -7.4540
# # s01 NC P7 n -5.6729
# # s01 NC P7 o -2.1812
# # s01 NC PO7 f -7.4169
#
#
# library(readr)
# library(dplyr)
#
# dati=lapply(datafiles, read_delim,col_names = FALSE ,delim = " ")
# dati=bind_rows(dati)
# str(dati)
# names(dati)=c("Subj","Group","Chan","Condition","Y")
#
# # Not used in this analysis
# dati$Group=NULL
# dati$Subj=factor(dati$Subj)
# dati$Chan=factor(dati$Chan)
# dati$Condition=factor(dati$Condition)
# str(dati)
# save(dati,file="datiEEG.Rdata")
#
# dati2=subset(dati,(Chan=="O1")&(Condition%in%c("f","n")))
# dati2$Condition=factor(dati2$Condition)
# save(dati2,file="dati2EEG.Rdata")
load("./dataset/datiEEG.Rdata")
load("./dataset/dati2EEG.Rdata")
# VERY IMPORTANT:
```

```
contrasts(datiChan) \leftarrow contr.sum(6)
contrasts(dati$Condition) <- contr.sum(5)
contrasts(dati$Subj) <- contr.sum(nlevels(dati$Subj))
```

```
contrasts(dati2$Condition) <- contr.sum(2)
contrasts(dati2$Subj) <- contr.sum(nlevels(dati2$Subj))
```
1.2 Motivation (EDA)

```
For Channel O1:
library(ggplot2)
p <- ggplot(subset(dati,Chan=="O1"),aes(Condition,Y))
p+geom_point(size = 3) +geom_boxplot(alpha=.1)
```


```
Is there a specificity of the subject?
```
dati01=subset(dati,Chan=="01") library(ggplot2) p <- ggplot(dati01,aes(Condition,Y)) $p+geom_point(aes(group = Subj, colour = Subj))+$ $\frac{1}{\text{geom_line}}(\text{aes}(\text{group} = \text{Subj}, \text{colour} = \text{Subj}))$ + geom_boxplot(alpha=.1)

We subtract the subject-specific effect (i.e. subject's mean) to each observation.

```
dati01=subset(dati,Chan=="01")
Y=scale(matrix(dati01$Y,5), scale=FALSE)
dati01$Y=as.vector(Y)
library(ggplot2)
p <- ggplot(dati01,aes(Condition,Y))
p+geom_point(aes(group = Subj, colour = Subj))+
```
geom_boxplot(alpha=.1)

geom_line(aes(group = Subj, colour = Subj))+

The dispersion of the data has been largely reduced. This effect is the one taken in account by the models for repeated measures.

2 Two Paired samples and symmetry test

2.1 Definition

Let consider the reduced problem: channel Chan=="O1 and Condition=="n" or Condition=="f".

Let be *y* the outcome, *x* the condition/treatment $(x \in \{1 = "f", 2 = "n"\}$ in our case). Let be $z =$ Subject the nuisance factor in the stratified problem.

Under the null hypothesis: $f(y|x=1, z) = f(y|x=2, z) = f(y|z)$ while possibly (even under H_0) $\exists (z, z') : f(y|x, z) \neq f(y|x', z')$

This imply that observations are exchangeable only within the same subject (*z*, i.e. playing the role of Strata).

In the gaussian-parametric approach we assume a different mean for each subject (i.e. Subject-specific effect), but the variance is forced to be constant among subjects. Here we don't make this assumption. This is much more realistic (see discussion later).

2.2 Testing Symmetry

The test statistic is based on the mean difference $T(y) = \sum_{i=1}^{n} (y_{i2} - y_{i1})/n = \sum_{i=1}^{n} d_i/n$ Since $f(y_{i1}) = f(y|x = 1, z = i) = f(y|x = 2, z = i) = f(y_{i2}), d_i$ is symmetric.

Therefore, the null hypothesis is equivalentely written as:

$$
H_0: f(y_i|x_i = 1, z_i) = f(y_i|x_i = 2, z_i) \,\forall z_i
$$

\n
$$
\implies f(d_i) = f(-d_i) \,\forall z_i
$$

Permutantion within the observation *i* reduces to randomly flipping the sign of *dⁱ* . We test for symmetry.

REMARK The opposite implication is not always true: $f(d_i) = f(-d_i) \Rightarrow f(y_i|x_i = 1, z_i) = f(y_i|x_i = 1, z_i)$ 2, z_i) ∀ z_i ; consider the important example of $y_i \sim N(0, \Sigma(x_i, z_i))$ (i.e. the variance depends on the levels of x_i and the subject z_i). In this case $(y_i|x_i = 1, z_i) - (y_i|x_i = 2, z_i) \sim N(0, \Sigma(x_i = 1, z_i) + \Sigma(x_i = 2, z_i))$ is still (normal and therefore) symmetric! Therefore the assumption of symmetry of the difference is broader than the assumption of exchangeability of observations within the same subject.

2.3 A bit of theory

(see also Pesarin, 2001; Hemerik & Goeman, 2017)

Let *Y* be data taking values in a sample space *Y*. Let Π be a finite set of transformations $\pi : \mathcal{Y} \to \mathcal{Y}$, such that Π is a group with respect to the operation of composition of transformations:

- it contains identity,
- every element has an inverse in the group,
- closure: if $\pi_1, \pi_2 \in \Pi$: $\pi_1 \circ \pi_2 \in \Pi$

(e.g. Π set of all possible permutations)

Test statistic $T(Y) : \mathbb{R}^n \to \mathbb{R}$

*H*₀: null hypothesis which implies that the joint distribution of the test statistics $T(\pi Y)$, $\pi \in \Pi$, is invariant under all transformations in Π of *Y*. That is, writing $\Pi = {\pi_1, \ldots, \pi_{|\Pi|}}$, under H_0 :

$$
T(\pi_1 Y), \ldots, T(\pi_{|\Pi|} Y) \stackrel{d}{=} T(\pi_1 gY), \ldots, T(\pi_{|\Pi|} gY)
$$

for all $q \in \Pi$.

Note that it holds when for all $\pi \in \Pi$: $Y \stackrel{d}{=} \pi Y$.

Orbit of O:

.

$$
\mathcal{O} = \{\pi Y : \pi \in \Pi\} \subseteq \mathcal{Y}
$$

(losely) the set of all samples having the same likelihood under H_0 .

$$
\mathcal{O} = \{ \pi \mathbf{y} : f(\pi \mathbf{y}) = f(\mathbf{y}) \}
$$

 $(|\mathcal{O}|)$ number of elements of \mathcal{O})

If we assume exchangeability of observations, then:

 $\mathcal{O} = \{ \text{all permutations of the observed data } \mathbf{y} \} = \{ \mathbf{y}^* : \pi^* \circ \mathbf{y} \}$

Remark: For Repeated Measures this means that, Under the Null Hypothesis, observations within subject are assumed to be exchangeable: $f(y_1, y_2) = f(y_2, y_1)$.

This assumption is always true as long as observations:

- are **identically distributed** (within the same subject),
- have the **same dependence**, e.g. the same correlation.

t-test and linear models assumes independence, which is just a special case: $(f(y_1, y_2) = f(y_2)f(y_1) =$ $f(y_2, y_1)$, i.e. a more severe assumption!

paired *t*-test and repeated measures assumes homoscedasticity of the vector of differences (i.e. one difference for each subject), here we only assume symmetry. Again, permutation approach makes less assumptions.

 $T^{(k)}(Y)$ $\lceil (1-\alpha)|\Pi| \rceil$ -th sorted value of $T(\pi Y)$

Theorem: Under H_0 , $P(T(Y) > T^{(k)}) \leq \alpha$.

intuition:

$$
f(\mathbf{y}^*|\mathcal{O}) = \frac{f(\mathbf{y}^* \cap \mathcal{O})}{f(\mathcal{O})} = \frac{f(\mathbf{y}^*)}{f(\mathcal{O})} = \frac{f(\mathbf{y}^*)}{f(\cup_{y \in \mathcal{O}} y)} = \frac{1}{|\mathcal{O}|} \ \forall \ \mathbf{y}^* \in \mathcal{O}
$$

i.e. each permutation is equally likely in the Orbit \mathcal{O} .

(due to group structure)

$$
P(T(\mathbf{y}) \ge T^{(k)} | \mathbf{y}^* \in \mathcal{O}, H_0) =
$$

=
$$
\int_{T^{(k)}}^{+\infty} f(T(\mathbf{y})) dT(\mathbf{y}) =
$$

=
$$
\sum_{\mathbf{y} \in \mathcal{O}} I(T(\mathbf{y}^*) \ge T(\mathbf{y})) / |\mathcal{O}| \le \alpha \quad \forall \mathcal{O}
$$

2.3.1 Properties (see Pesarin, 2001)

The theorem above proves that the permutation tests have **exact control of the type I error**, i.e. $P(p - p)$ *value* $\leq \alpha | H_0$ = α assuming $\alpha \in \{1/|\mathcal{O}|, 2/|\mathcal{O}|, \dots, 1\}$ - don't forget that the orbit $\mathcal O$ is a finite set; if this is not the case, the test is (slightly) conservative.

Further properties:

- The permutations tests are **Unbiased**: $P(p value \le \alpha | H_0) > \alpha$
- The test is **Consistent**: $P(p value \le \alpha | H_1) \rightarrow 1$ when $n \rightarrow \infty$
- The test converge to the parametric counterpart (when it exists)

2.4 Results

The parametric paired t-test:

```
t.test(dati2$Y[dati2$Condition=="n"]-
         dati2$Y[dati2$Condition=="f"])
##
## One Sample t-test
##
## data: dati2$Y[dati2$Condition == "n"] - dati2$Y[dati2$Condition == "f"]
## t = 3.287, df = 19, p-value = 0.003877
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 0.2303616 1.0380084
## sample estimates:
## mean of x
## 0.634185
The nonparametric one:
```

```
flip(dati2$Y[dati2$Condition=="n"]-
        dati2$Y[dati2$Condition=="f"],perms=10000)
##
## Test Stat tail p-value
## Y t 3.287 >< 0.0038
# Equivalent to
flip(Y~Condition,Strata=~Subj,data=dati2,perms=10000)
##
## Test Stat tail p-value
## Y t 3.287 >< 0.0038
```
3 Permutation-Based Repeated Measures ANOVA

3.1 Introduction

Let be *y* the outcome, *x* the condition/treatment $(x \in (f, h, d, n, o)$ in our case). Let be $z =$ Subject the nuisance factor in a stratified problem.

Under the null hypothesis: $f(y|x, z) = f(y|x', z) = f(y|z) \; \forall (x, x')$ while possibly (even under H_0) $\exists (z, z') : f(y|x, z) \neq f(y|x', z')$

This imply that observations are exchangeable only within the same subject (*z*).

In the gaussian-parametric approach we assume a different mean for each subject (i.e. Subject-specific effect), but the variance is forced to be constant among subjects. Here we don't make this assumption. This is much more realistic (see discussion later).

3.2 One Linear model for each Subject

The following approach is developed in Basso & Finos (2012) and in Finos & Basso (2014). It is very similar to (basically an extension of) the NonParametric Group level analysis developed in SnPM by Nichols and Holmes (2001).

Assuming a hierarchical linear model we have a subject-specific linear model for each subject. The vector of outcomes *yⁱ* can be modeled as a linear function of the predictor

 $y_i = \beta_{0i} + \beta_{1i}x_1 + \ldots + \beta_{ki}x_2 + \varepsilon_i = X\beta_i + \varepsilon_i; i = 1 \ldots, n$

 y_i is a vector, *X* is a matrix of *k* predictors, ε_i is a vector of iid r.v.

Therefore: $(\hat{\beta}_i | \beta_i) \sim (\beta_i, \Sigma_i)$

We also assume that $\beta_i \sim (\beta, \Psi)$. Therefore we have that: $\hat{\beta}_i \sim (\beta, \Psi + \Sigma_i)$

Let's consider again the test on the effect of the Condition. The matrix of predictors X is the matrix of $k = C - 1 = 4$ dummy variables. The null hypothesis becomes:

$$
H_0: \beta = (\beta_1, \ldots, \beta_k) = 0
$$

The observations within the subject (i.e. stratum) are exchangeable under the null hypothesis. However, we don't have a *separate* (in the sense of Pesarin, 2001. This is often known as Subset Pivotality, WY, 1993) test for $H_0: \beta_1 = 0$ since we are permuting the observation under the global null hypothesis (i.e. values relative to other conditions affect the estimate of β_1).

The same problem rises if the matrix of *X* contains other factors, such as ChanL and Lateral. Let's now consider the experimental design modeled as a linear model with ChanL*Lateral*Condition.

Data preparation:

```
dati$Lateral=dati$Chan
levels(dati$Lateral)
## [1] "O1" "O2" "P7" "P8" "PO7" "PO8"
levels(dati$Lateral)[c(1,3,5)]="Left"
levels(dati$Lateral)[-1]="Right"
levels(dati$Lateral)
## [1] "Left" "Right"
dati$ChanL=dati$Chan
# https://en.wikipedia.org/wiki/Regular_expression
# Digits: \d
levels(dati$ChanL)=gsub("\\d","",levels(dati$ChanL))
```

```
contrasts(dati$Condition)=NULL
contrasts(dati$ChanL)=NULL
contrasts(dati$Lateral)=NULL
```
We fit a model *within* each subject and we store the vector of estimated coefficients.

dataCoeff=obs2coeffWithin(Y~ ChanL*Lateral*Condition,data=dati,units=~Subj)

From the results above we have that $\hat{\beta}_i \sim (\beta, \Psi + \Sigma_i)$. Since the model is full-rank in this design (i.e. no residuals error) and balanced, the distribution is also symmetric around the mean: $f(\hat{\beta}_i - \beta) = f(-(\hat{\beta}_i - \beta))$ (at least component-wise, i.e. for marginal test)

We know, then, how to derive an exact test for H_0 : $\beta = 0$. Now we use these coefficients as a dataset to test for multivariate symmetry.

A very important result relies in the fact that the tests are separate, that is, the symmetry under the null hypothesis is ensured for each element of *β* whatever is the true mean of the others elements of *β*.

```
mod=flipMixWithin(data=dataCoeff,perms=10000,statTest = "Tnaive")
summary(mod)
```

```
## Call:
## flipMixWithin(perms = 10000, data = dataCoeff, statTest = "Tnaive")
## 9999 permutations.
##
## est.Su Test Stat tail p-value sig.
## (Intercept)_Tnaive NA Tnaive -3.1020 >< 0.0058 **
## ChanLP_Tnaive NA Tnaive -7.6164 >< 0.0002 ***
## ChanLPO_Tnaive NA Tnaive -5.2511 >< 0.0008 ***
## LateralRight_Tnaive NA Tnaive -1.7162 >< 0.0996
## Conditionf_Tnaive NA Tnaive 0.9595 >< 0.3456
## Conditionh_Tnaive NA Tnaive 4.8141 >< 0.0008 ***
## Conditionn_Tnaive NA Tnaive 4.5430 >< 0.0002 ***
## Conditiono_Tnaive NA Tnaive 6.7913 >< 0.0002 ***
## ChanLP:LateralRight_Tnaive NA Tnaive 0.5942 >< 0.5588
## ChanLPO:LateralRight_Tnaive NA Tnaive 2.4635 >< 0.0234 *
## ChanLP:Conditionf_Tnaive NA Tnaive -0.3815 >< 0.7118
## ChanLPO:Conditionf_Tnaive NA Tnaive -0.7127 >< 0.4800
## ChanLP:Conditionh_Tnaive NA Tnaive -1.2310 >< 0.2270
## ChanLPO:Conditionh_Tnaive NA Tnaive -1.0945 >< 0.2850
## ChanLP:Conditionn_Tnaive NA Tnaive 0.1750 >< 0.8640
```


Since the fitted model within each subject is a saturated one, the effect of each nuisance factor is removed without errors in the estimates.

p-values for Global effect and factor are computed by NPC methodology.

```
#Global
npc(mod,trace=FALSE)
##
## comb.funct nVar Stat p-value
## V1 Fisher 30 86.83 0.0002
# colnames(mod@permT)
# Combined by factor
summary(npc(mod,subsets = list(ChanL=1:2,Lateral=3,Condition=4:7,
                  ChanL Lateral=8:9,
                  ChanL_Condition=10:17,
                  Lateral_Condition=18:21,
                  ChanL_Lateral_Condition=22:29)))
## Call:
# npc(permTP = mod, subsets = list(ChanL = 1:2, Lateral = 3, Condition = 4:7, ChanL_Lateral = 8:
## permutations.
##
## comb.funct nVar Stat p-value sig.
## ChanL Fisher 2 13.667 0.0002 ***
## Lateral Fisher 1 7.131 0.0008 ***
## Condition Fisher 4 19.017 0.0004 ***
## ChanL_Lateral Fisher 2 9.099 0.0004 ***
## ChanL_Condition Fisher 8 16.464 0.0446 *
## Lateral_Condition Fisher 4 9.909 0.0224 *
## ChanL_Lateral_Condition Fisher 8 10.416 0.2074
```
This approach is very general, it works also when we have non-balanced and non-saturated models. For example we may like to use the values computed on each single trial and not the average. In this case, the model is not full rank and we have different number of trial in each condition. In this case, the test remains exact only if we can assume symmetric distribution of the errors. If this is not the case, the test become approximated. It should be noted, however that usually the speed in convergence to the exact control of the type I error is faster than for parametric (gaussian) model since the our method becomes exact whenever the distributions of the $\hat{\beta}_i$ become symmetric (i.e. odd moments converge to zero), while the parametric one is exact when the same distributions become normal (i.e. both even and odd moments converge to zero).

Furthermore, the symmetry is the only condition, we still allow for hetheroscedastic errors between subjects. We also allow for different design matrix between subjects $(X_i \neq X_j)$. As an extreme and surprisingly result, we also allow for the error within subject to varies.

3.3 Comparison with parametric Mixed-Model

Compare the previous results with the one of the parametric approach:

```
library(lmerTest)
## Caricamento del pacchetto richiesto: lme4
## Warning: il pacchetto 'lme4' è stato creato con R versione 4.1.2
## Caricamento del pacchetto richiesto: Matrix
##
## Caricamento pacchetto: 'lmerTest'
## Il seguente oggetto è mascherato da 'package:lme4':
##
## lmer
## Il seguente oggetto è mascherato da 'package:stats':
##
## step
mod_mix=lmer(Y~ ChanL*Lateral*Condition +(ChanL*Lateral|Subj),data=dati)
car::Anova(mod_mix)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: Y
## Chisq Df Pr(>Chisq)
## ChanL 81.2681 2 < 2.2e-16 ***
## Lateral 6.0196 1 0.01415 *
## Condition 1200.8786 4 < 2.2e-16 ***
## ChanL:Lateral 7.7555 2 0.02070 *
## ChanL:Condition 74.4535 8 6.346e-13 ***
## Lateral:Condition 0.8502 4 0.93160
## ChanL:Lateral:Condition 2.8455 8 0.94367
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
The (parametric) Repeated measures method assumes of the errors, the nonparametric (permutation) does not.

Mixed model does not make this assumption, but assumes normality and homoscedasticity of the observations.

This could be a relevant point for our data:

```
dati$residuals=residuals(mod_mix)
p <- ggplot(dati, aes(x=Subj, y=residuals,fill=Condition)) + geom_boxplot()
p
```


The variability of the subjects doesn't not appear homogeneous.

3.4 Contrasts and post-hoc

3.4.1 Post-hoc and Custom contrasts

For this example we restrict the analysis to the comparison Happy vs Neutral.

Let' now compute the vectors of contrasts (one vector of reach channel, length equal to number of subjects): Happy vs Neutral

```
contrasts(dati$Chan)=NULL
contrasts(dati$Condition)=NULL
dati_2cond_6chan=subset(dati,(Condition%in%c("h","n")))
dati_2cond_6chan$Condition=factor(dati_2cond_6chan$Condition)
dati_2cond_6chan$Condition=relevel(dati_2cond_6chan$Condition,"n")
```

```
dataCoeff=obs2coeffWithin(Y~ Chan*Condition,data=dati_2cond_6chan,units=~Subj)
```
colnames(dataCoeff\$coeffWithin)

dataCoeff\$coeffWithin contains the estimates of the coefficients for the model for each subject.

The column Conditionh contains the contrast $h-n$ in Chan01. To get the same contrast in any other channel we add the column of interaction Conditionh: Chan to the one of Chan01. As an example, to get n-f in Chan02 we make Conditionh + ChanO2:Conditionh.

```
Y=dataCoeff$coeffWithin[,7]+cbind(ChanO1=0,dataCoeff$coeffWithin[,8:12])
colnames(Y)=gsub(":Condition.","",colnames(Y))
colnames(Y)
```
[1] "ChanO1" "ChanO2" "ChanP7" "ChanP8" "ChanPO7" "ChanPO8"

```
Analysis: raw and adjusted p-values (min-p procedure)
```

```
res=flip(Y,perms=10000)
```

```
res=flip.adjust(res)
summary(res)
```

```
## Call:
## flip(Y = Y, \text{perms} = 10000)## 9999 permutations.
##
## Test Stat tail p-value Adjust:maxT sig.
## ChanO1 t -1.472 >< 0.1594 0.1594
## ChanO2 t -2.137 >< 0.0438 0.0744
## ChanP7 t -2.449 >< 0.0264 0.0588
## ChanP8 t -3.675 >< 0.0022 0.0056 **
## ChanPO7 t -2.375 >< 0.0322 0.0620
## ChanPO8 t -2.981 >< 0.0060 0.0212 *
```
To show the results plot intensity colors based on the significance of the adjusted p-values for each channel. Just for visual purposes, we transform the adjusted p-values by −*log*10(*p*):

```
# install.packages("legit")
library(eegkit)
```
Warning: il pacchetto 'eegkit' è stato creato con R versione 4.1.3

Warning: il pacchetto 'rgl' è stato creato con R versione 4.1.2

```
# get the a z value from the adjusted p-value for each channel, just for visual purposes:
pvals=getFlip(res,"Adjust:maxT")
rownames(pvals)=gsub("Chan","",rownames(pvals))
```
match to eeg coordinates data(eegcoord) cidx <- match(rownames(pvals),rownames(eegcoord))

```
# plot t-stat in 2d
eegspace(eegcoord[cidx,4:5],-log10(pvals[,1]),cex.point = 3,colorlab="-log10(adj-p)",mycolors=heat.colors
```


4 (minimal) Bibliography

- Basso & Finos (2012). Exact Multivariate Permutation Tests for Fixed Effects in Mixed-Models. *Communications in Statistics - Theory and Methods*, 41: 2991 - 3001.
- Finos & Basso (2014) Permutation tests for between-unit fixed effects in multivariate generalized linear mixed models. *Stat Comput* 24, 941–952.
- Nichols & Holmes (2001) Nonparametric Permutation Tests For Functional Neuroimaging: A Primer with Examples. *Human Brain Mapping* 15, 1-25.<https://onlinelibrary.wiley.com/doi/ftr/10.1002/hbm.1058>