R\_HW\_8\_Repeated Measures

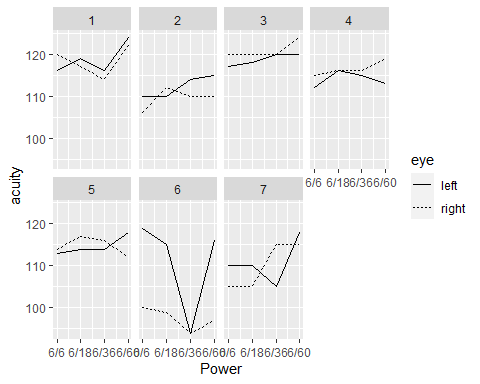
Eom SangJun

2020 11 19

7명의 피실험자들을 대상으로 시각의 예민함에 대해 측정한 자료를 통해 반복측정 자료에 대한 분석을 해보자. 이 데이터에서 종속변수는 눈의 피질에서 광선 플래시와 반응 사이의 지연시간(밀리초 단위)을 의미한다. 그리고 양 안은 네 개의 서로 다른 도수(power)의 렌즈에 대해 각각 한 번씩 실험 되었다.

data(vision, package='faraway')

분석하기에 앞서 도수가 증가함에 따라 acuity가 어떻게 변하는 지를 확인하기 위해 power에 관한 새로운 numerical변수를 만들어주었다.   
vision$npower <- rep(1:4, 14)  
ggplot(vision, aes(y=acuity, x=npower, linetype=eye)) +  
 geom\_line() +  
 facet\_wrap(~ subject, ncol=4) +  
 scale\_x\_continuous('Power', breaks = 1:4, label=c('6/6', '6/18', '6/36', '6/60'))



* 어떠한 경향성이나 오른쪽 눈과 왼쪽 눈의 차이는 잘 드러나지 않는 것으로 보인다.
* 단, 6번 피험자의 경우 양쪽 눈이 큰 차이를 보인다. 또한 이 실험자의 세 번째 측정은 에러가 있는 것으로 보인다.

이제 어떻게 모델링을 할 것인지를 정해보자.

우선 power는 fixed 데이터이다. 우선은 nominal factor로 다룰 것이지만 정량적인 방법으로도 fitting할 수 있다. 피실험자의 경우 큰 모집단에서 랜덤으로 뽑은 것이라고 가정하는 것이 적절할 것이기 때문에 random effect로 볼 수 있다. 또한 우리는 피험자들 간에 오른쪽 눈과 왼쪽 눈의 일관된 차이가 없다고 생각하기 때문에 eye factor가 subject 안에서 nested 되어있다고 가정한다.

위의 가정사항들을 토대로 model을 세워보면 다음과 같다. (만약 우리가 consistent한 left vs right eye effect가 있다고 믿는다면 우리는 eye를 fixed effect로 처리해야 한다. 즉, (1|subject:eye) 자리에 eye를 넣어야 한다.

mmod <- lmer(acuity ~ power + (1|subject) + (1|subject:eye), vision)

위의 식을 표현하면,

Summary를 살펴보자.  
sumary(mmod)

## Fixed Effects:  
## coef.est coef.se  
## (Intercept) 112.64 2.23   
## power6/18 0.79 1.54   
## power6/36 -1.00 1.54   
## power6/60 3.29 1.54   
##   
## Random Effects:  
## Groups Name Std.Dev.  
## subject:eye (Intercept) 3.21   
## subject (Intercept) 4.64   
## Residual 4.07   
## ---  
## number of obs: 56, groups: subject:eye, 14; subject, 7  
## AIC = 342.7, DIC = 349.6  
## deviance = 339.2

피험자에 대한 estimated standard deviation은 4.64이며 피험자가 주어졌을 때 eye에 대한 것은 3.21이다. Residual의 경우 4.07이다.

우리가 여기서 사용한 random effect의 구조는 동일한 피험자에 대한 측정 간 correlation과 동일한 eye에 대한 측정 간의 correlation을 유도한다. 이 두 correlation을 직접 계산해보면 다음과 같다.

4.64^2/(4.64^2+3.21^2+4.07^2)

## [1] 0.4448393

🡪 동일한 피험자에 대한 측정들 간 correlation

(4.64^2+3.21^2)/(4.64^2+3.21^2+4.07^2)

## [1] 0.6577401

🡪 동일한 eye에 대한 측정들 간의 correlation

🡪 우리가 예상한대로, 동일한 피험자의 양 안에 대한 correlation보다는 다른 피험자들 간의 동일한 눈(오른쪽이든 왼쪽이든)에 대한 correlation이 더 높다는 것을 알 수 있다.

이번에는 Kenward-Roger adjusted F-test를 이용하여 power effect를 확인해보자.

mmod <- lmer(acuity ~ power + (1|subject) + (1|subject:eye), vision, REML = FALSE)  
nmod <- lmer(acuity ~ 1 + (1|subject) + (1|subject:eye), vision, REML=FALSE)  
KRmodcomp(mmod, nmod)

## F-test with Kenward-Roger approximation; time: 0.14 sec  
## large : acuity ~ power + (1 | subject) + (1 | subject:eye)  
## small : acuity ~ 1 + (1 | subject) + (1 | subject:eye)  
## stat ndf ddf F.scaling p.value   
## Ftest 2.8263 3.0000 39.0000 1 0.05106 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

🡪 우리의 예상과는 다르게 power effect가 significant하지 않다고 나온다. 그런데 우리는 앞서 6번째 피험자의 데이터 중 하나가 에러가 있는 것 같다는 것을 발견했다. 따라서 이를 제거해보고 다시 확인해보자.

mmodr <- lmer(acuity ~ power + (1|subject) + (1|subject:eye), vision, REML=FALSE, subset=-43)  
nmodr <- lmer(acuity ~ 1 + (1|subject) + (1|subject:eye), vision, REML=FALSE, subset = -43)  
KRmodcomp(mmodr, nmodr)

## F-test with Kenward-Roger approximation; time: 0.07 sec  
## large : acuity ~ power + (1 | subject) + (1 | subject:eye)  
## small : acuity ~ 1 + (1 | subject) + (1 | subject:eye)  
## stat ndf ddf F.scaling p.value   
## Ftest 3.5956 3.0000 38.0373 1 0.02212 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

🡪 제거해보고 확인해보니 significant하다는 결론을 얻을 수 있었다.

그런데 이는 오직 제일 높은 도수에서의 effect 때문에 나타난 것이다. 우리는 Helmert contrasts를 통해서 가장 높은 도수가 다른 세 level의 평균보다 acuity가 더 높게 나타난다는 것을 확인할 수 있다.

op <- options(contrasts = c('contr.helmert', 'contr.poly'))  
mmodr <- lmer(acuity ~ power + (1|subject) + (1|subject:eye), vision, subset=-43)  
  
sumary(mmodr)

## Fixed Effects:  
## coef.est coef.se  
## (Intercept) 113.79 1.76   
## power1 0.39 0.54   
## power2 0.04 0.32   
## power3 0.71 0.22

🡪 effect sizes 대비 standard error를 보면, 우리는 오직 세 번째 contrasts만 significant하다는 것을 확인할 수 있다.

##   
## Random Effects:  
## Groups Name Std.Dev.  
## subject:eye (Intercept) 4.97   
## subject (Intercept) 2.87   
## Residual 2.88   
## ---  
## number of obs: 55, groups: subject:eye, 14; subject, 7  
## AIC = 319.2, DIC = 308.9  
## deviance = 307.1

options(op)

Helmert contrast matrix는 다음과 같다.  
contr.helmert(4)

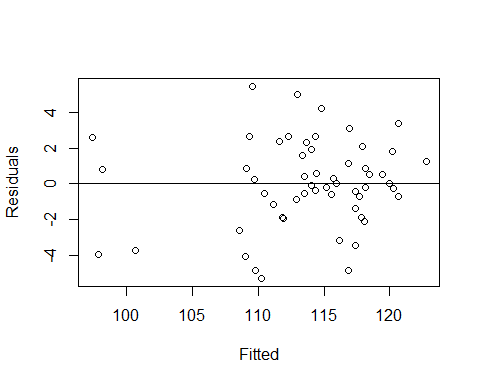
## [,1] [,2] [,3]  
## 1 -1 -1 -1  
## 2 1 -1 -1  
## 3 0 2 -1  
## 4 0 0 3

🡪 세 번째 contrast column은 네 번째 level과 나머지 세 level의 평균과의 차이를 나타낸다. 마찬가지로 나머지 두 contrast에 비해 significant한 것을 볼 수 있다.

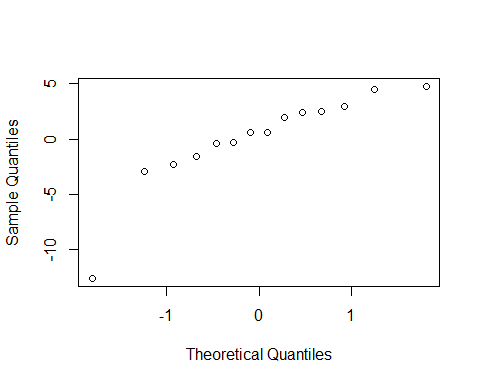
마지막으로 diagnostic plot을 살펴보자.

Residual and fitted values와 QQ plot of random effect for the eyes는 다음과 같다.

plot(resid(mmodr) ~fitted(mmodr), xlab='Fitted', ylab='Residuals')  
abline(h=0)



qqnorm(ranef(mmodr)$'subject:eye'[[1]], main='')



* 6번째 피험자의 오른쪽 눈에 대한 데이터가 outlier로 나타난다는 것을 알 수 있다. 추가적인 analysis를 위해서 subject 6의 데이터를 버리는 것을 고려할 수 있다. 데이터 수 자체가 워낙 적어서 data의 손실이 아쉽기는 하지만 피하기 어려워 보인다. 결국 더 확실한 결론을 위해서는 추가적인 데이터 수집이 필요해 보인다.