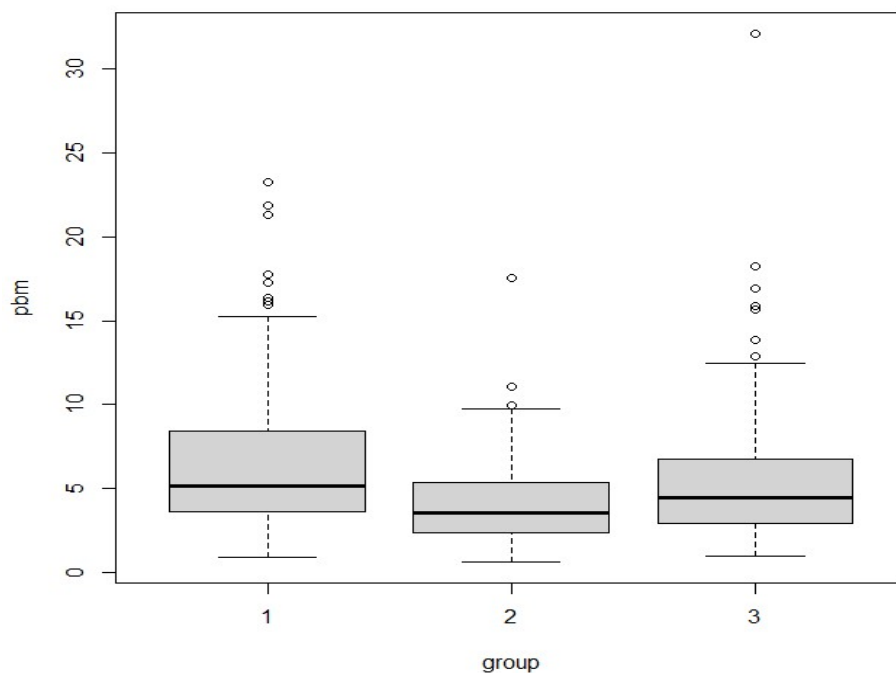


Question 1:

In this nested design, fixed factor is 'group' and random factors is 'patient'. The response of this experiment is 'pbm'.

Question 2(a):



Question 2(b):

According to the boxplot in 2(a), apparently, there is not a significant difference between the groups, they have similar median around 5. All groups have some outliers and especially group 1&3 have many outliers which may affect the outcome. More analyses need to be conducted.

Question 2(c):

To fit the model, we need R code such as:

“pbm ~ group + Error(group : patient), data = dat”

Because R has the rule, fixed term must be apart from error if exist.

Question 2(d):

```
> dat.aov1 = aov(pbm ~ group + Error(group:patient), data=dat)
aov(pbm ~ group + Error(group:patient), data = dat)에서 경고가 발생했습니다
:
  Error() 모델이 singular합니다
> summary(dat.aov1)
```

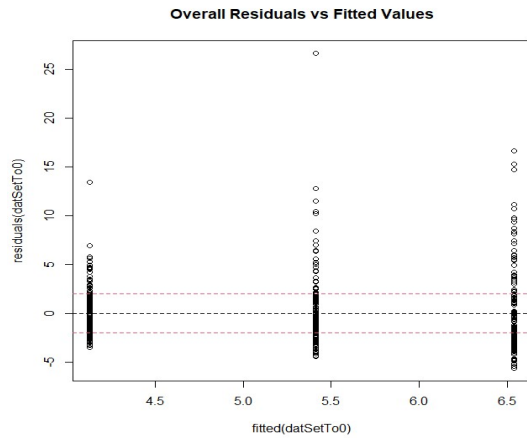
```
Error: group:patient
      Df Sum Sq Mean Sq F value    Pr(>F)
group    2  434.8   217.42   16.01 0.0000012 ***
Residuals 87 1181.7    13.58
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Error: within
      Df Sum Sq Mean Sq F value    Pr(>F)
Residuals 360  4970    13.81
```

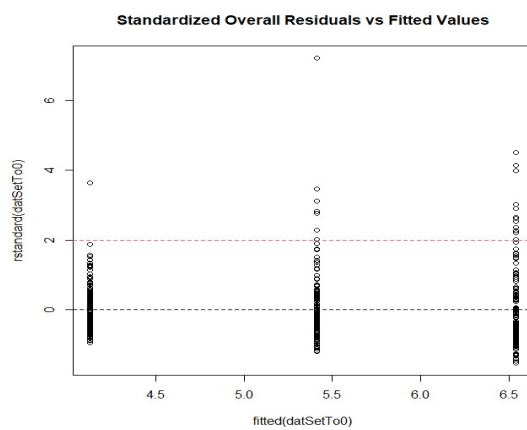
Question 2(e):

```
>>> datSetTo0 = lm(pbm ~ group, data = dat)
```

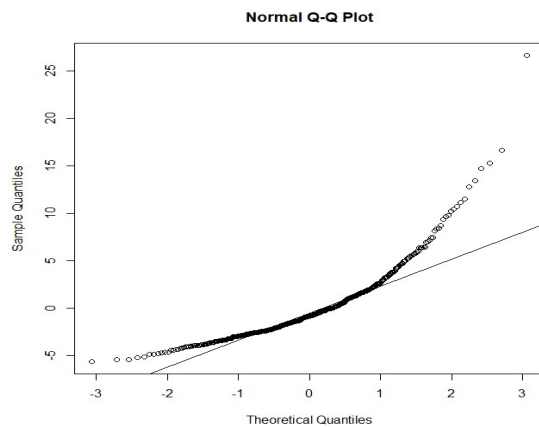
i. "Overall" Residuals vs Fitted Values



ii. Standardized "Overall" Residuals vs Fitted Values



iii. Q-Q plot for the Standardized Residuals



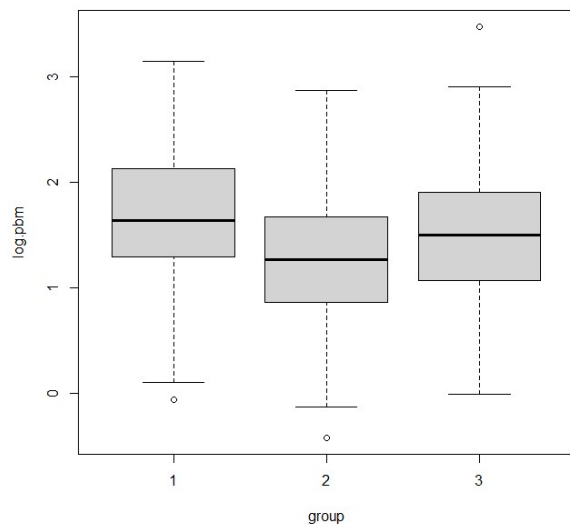
Question 2(f):

According to the three diagnostic plots in 2(e), it appears that the overall residuals look skewed to right from Q-Q plot. Presence of outliers and increasing trend in the residual plots. Therefore, this experiment did violate assumptions in ANOVA.

Question 2(g):

Based on the ANOVA table in 2(d), F-statistic of treatment is 16.01 and p-value is 0.0000012. It tells that group has significant impact on measurement of the perimeter of the airway basement membrane when statistically significant level is at 5% (α). However, it violates some assumptions and too many outliers exist which can lead incorrect analysis. Therefore, although, result shows effectiveness, more tests or better data are needed.

Question 3(a):



Question 3(b):

According to the boxplot in 3(a), compare to previous boxplot, most of outliers has been reduced and each quartile has decent range. There may be difference between median among groups, but more analyses need to be conducted.

Question 3(c):

To fit the model, we need R code such as:

“log.pbm ~ group + Error(group : patient), data = dat”

Question 3(d):

```
> dat.aov2 = aov(log.pbm ~ group + Error(group : patient), data = dat)
aov(log.pbm ~ group + Error(group:patient), data = dat)에서 경고가 발생했습니다 :
  Error() 모델이 singular합니다
> summary(dat.aov2)

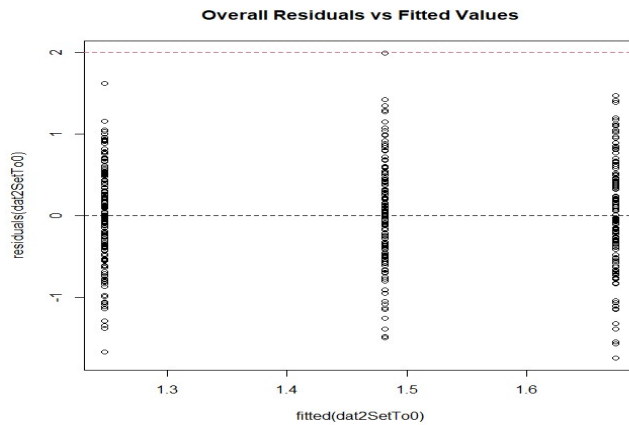
Error: group:patient
      Df Sum Sq Mean Sq F value    Pr(>F)
group    2  13.67   6.837   14.99 0.00000255 ***
Residuals 87  39.68   0.456
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: within
      Df Sum Sq Mean Sq F value Pr(>F)
Residuals 360  137.6   0.3822
```

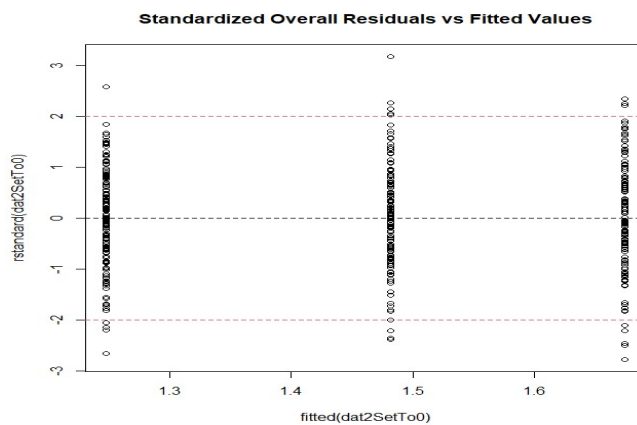
Question 3(e):

```
>>> dat2SetTo0 = lm(log.pbm ~ group, data = dat)
```

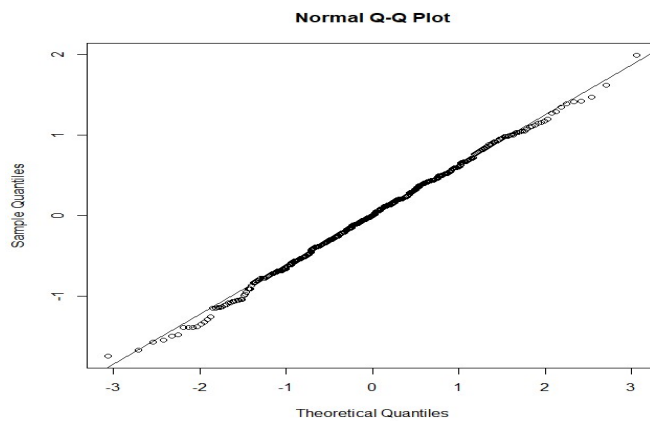
i. "Overall" Residuals vs Fitted Values



ii. Standardized "Overall" Residuals vs Fitted Values



iii. Q-Q plot for the Standardized Residuals



Question 3(f):

According to three diagnostic plots in 3(d), different to previous data's diagnostic, there appears to be no non-normality symptom from Q-Q plot nor heterogeneity from residuals plots. Therefore, this data did not violate any of assumptions in ANOVA.

Question 3(g):

Based on the ANOVA table in 3(d), F-statistic of treatment is 14.99 and p-value is 0.00000255. It tells that group has significant impact on measurement of the perimeter of the airway basement membrane. Since it did not violate any of assumptions, treatment(group) effect is statistically significant level is at 5% (α). However, to distinguish which group has effectiveness and how much it impacts, more statistical analyses need to be conducted.

Question 4(a):

```
> a = 3
> b = 30
> r = 5
> MS.Res = sum(residuals(dat.aov2$'group:patient')^2) / df.residual(dat.aov2$'group:patient')
> MS.Res
[1] 0.4561076
> SED = sqrt(2*MS.Res/(b*r))
> SED
[1] 0.07798355
```

SED = 0.07798355

Question 4(b):

```
> LSD = qt(1-0.05/2, a*(b-1)) * SED
> LSD
[1] 0.1550008
```

LSD = 0.1550008

Question 4(c):

```
> Means = with(dat, tapply(pbm, list(group), mean))
> D12 = Means[1] - Means[2]
> D23 = Means[2] - Means[3]
> D13 = Means[1] - Means[3]
> rbind(
+   c(D12-LSD, D12+LSD),
+   c(D13-LSD, D13+LSD),
+   c(D23-LSD, D23+LSD)
+ )
      1      1
[1,] 2.2511992 2.561201
[2,] 0.9702659 1.280267
[3,] -1.4359341 -1.125933
```

Question 5:

According to 4(c), none of group has the same effect on pbm. Because differences from group means greater than least significant difference. If groups have same impact, computed 95% confidence intervals for the difference of group means should include 0. Therefore, all the groups have statistically different effect on pbm.

According to 4(c), both 1-2 and 1-3 has positive value which means Group 1 has the highest mean and 2 – 3 is negative which shows Group 3 is bigger than Group 2. Therefore, Group 1 has the highest and Group 2 has the lowest outcome.