

lm assumptions

Guannan Shen

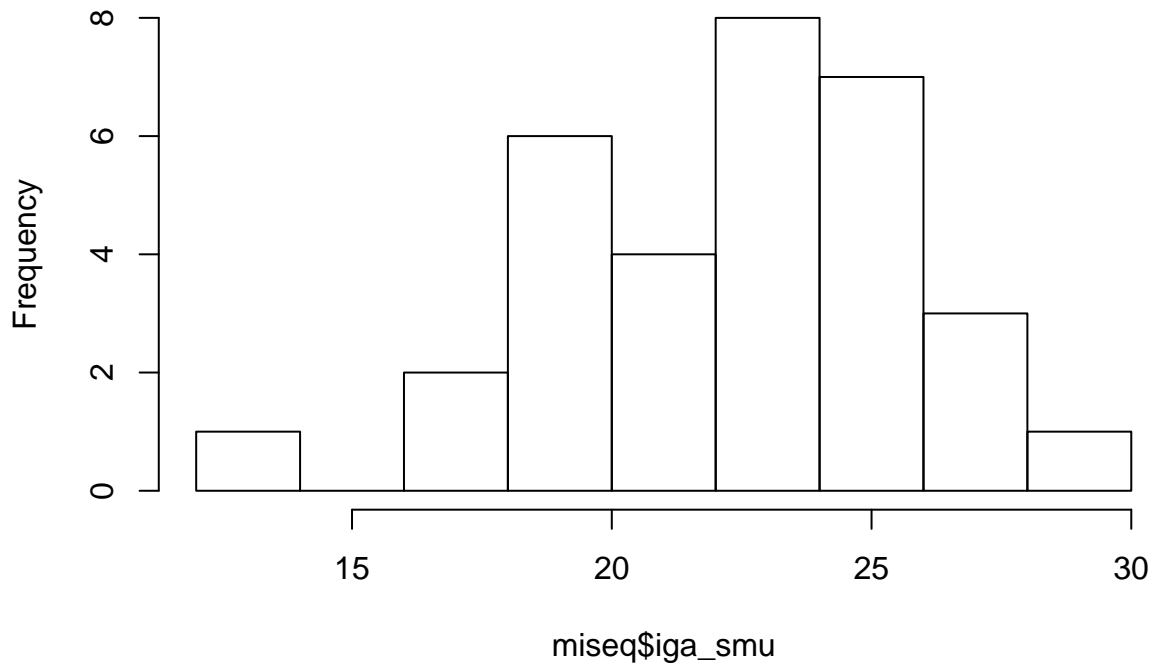
March 24, 2019

```
## [1] "/home/guanshim/Documents/gitlab/Cario_RNASeq_Microbiom_Inte/DataProcessed/Vgenes"
##### check the lm assumption of single mutation
#####
miseq <- read_excel("miseq.xlsx")
head(miseq)
```

```
## # A tibble: 6 x 5
##   ID      HIV   iga_smu igk_smu igl_smu
##   <chr>   <chr>   <dbl>   <dbl>   <dbl>
## 1 MIHIV138 no      24.3    10.9    18.1
## 2 MIHIV178 no      24.9    11.2    13.6
## 3 MIHIV255 no      23.2    11.3     15
## 4 MIHIV278 no      24.7    15.5    NA
## 5 MIHIV361 no      22.3     9.22   NA
## 6 MIHIV404 no      18.8    14.2    NA
```

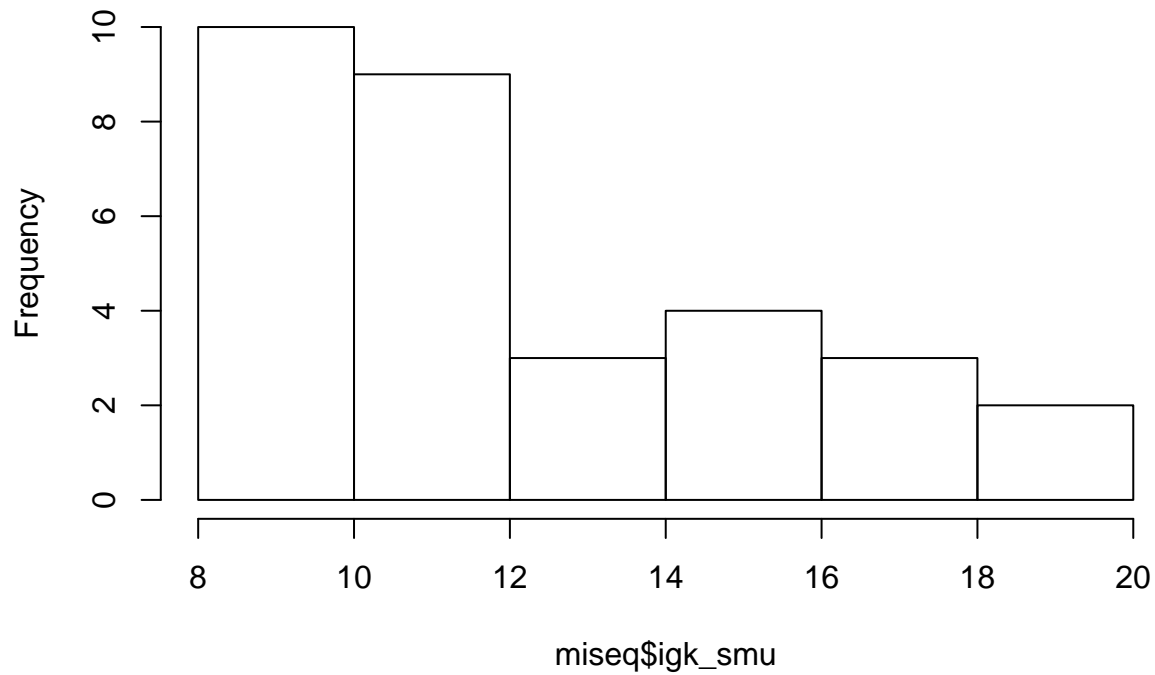
```
hist(miseq$iga_smu)
```

Histogram of miseq\$iga_smu



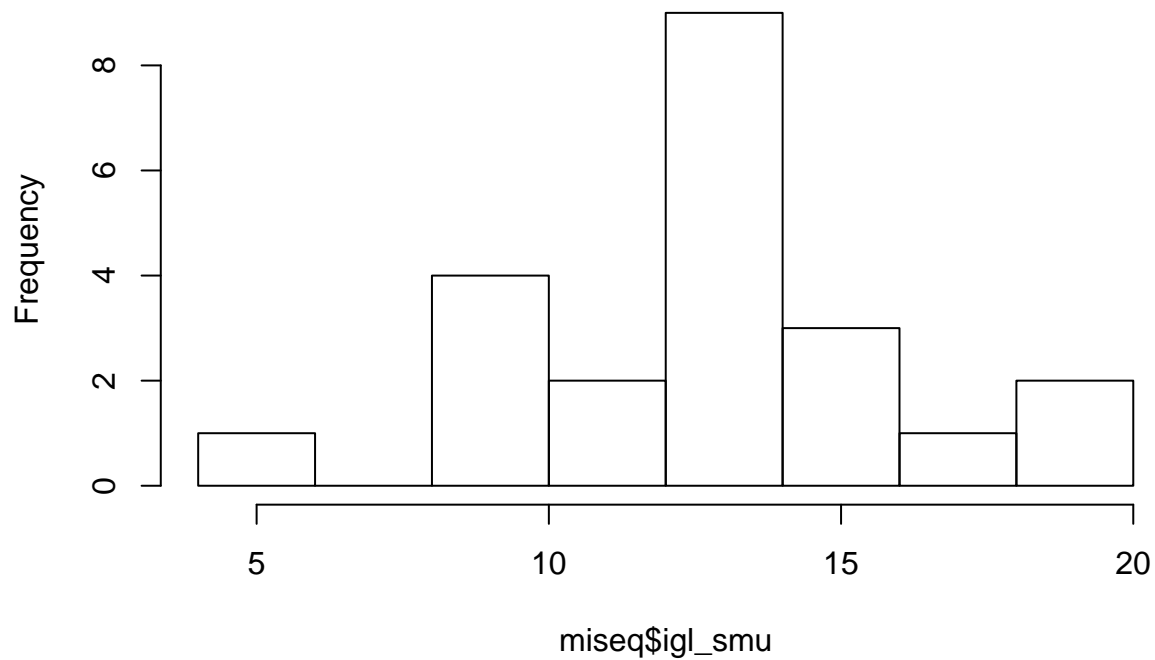
```
hist(miseq$igk_smu)
```

Histogram of miseq\$igk_smu



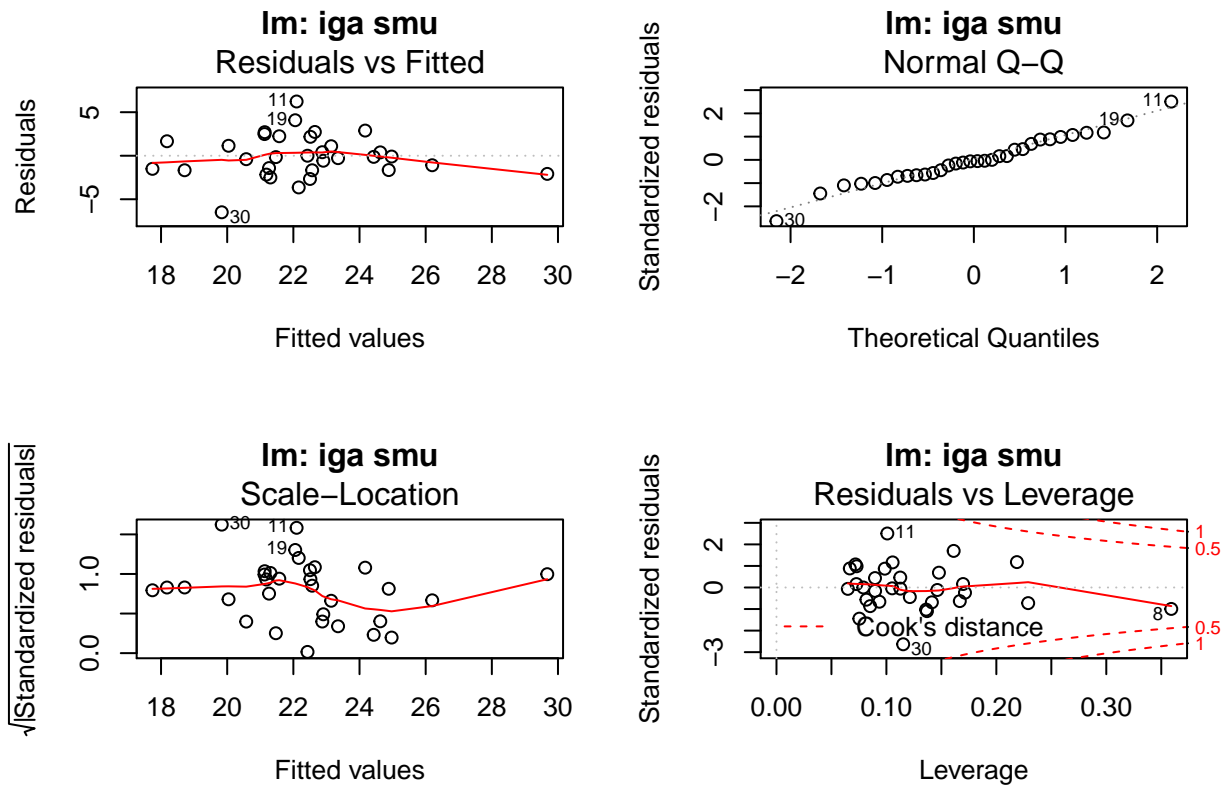
```
hist(miseq$igk_smu)
```

Histogram of miseq\$igl_smu

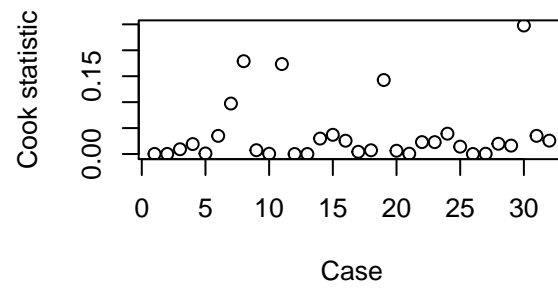
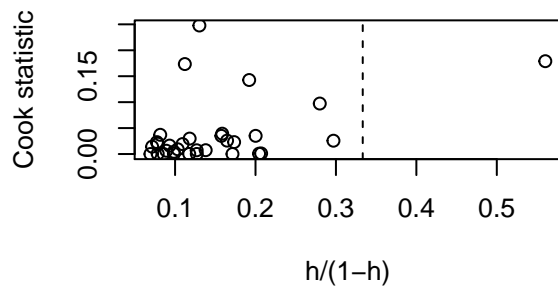
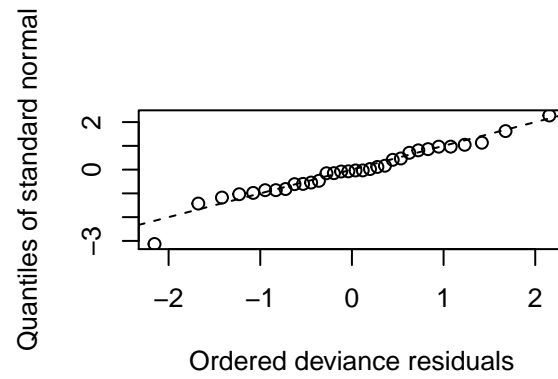
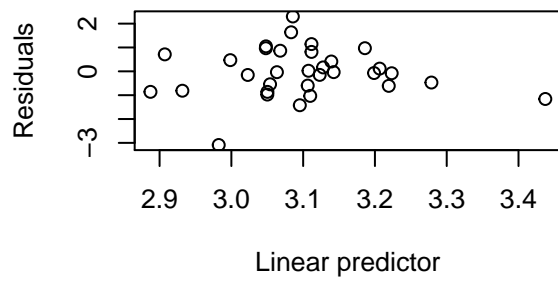


```
# diagnostic plots Xs  
Clin <- read_excel("clin_microbiome.xlsx")  
##### iga
```

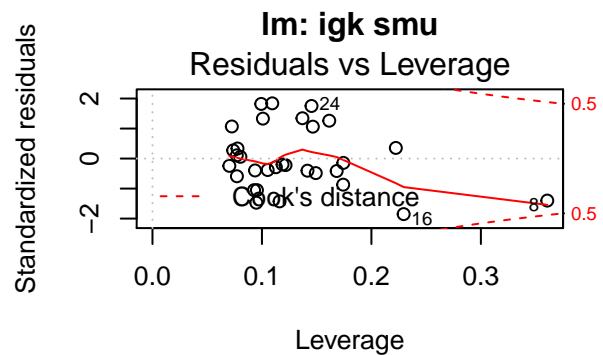
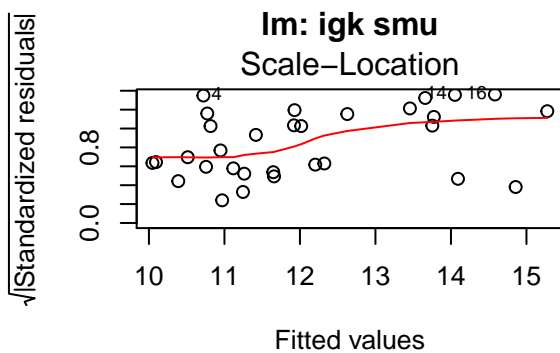
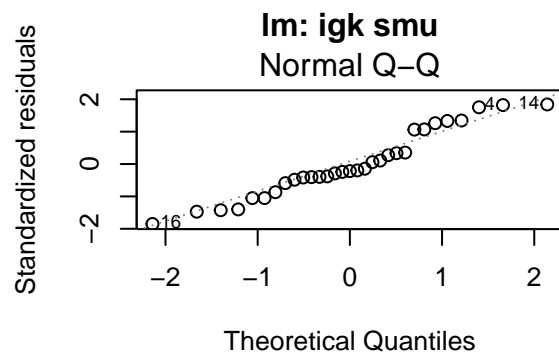
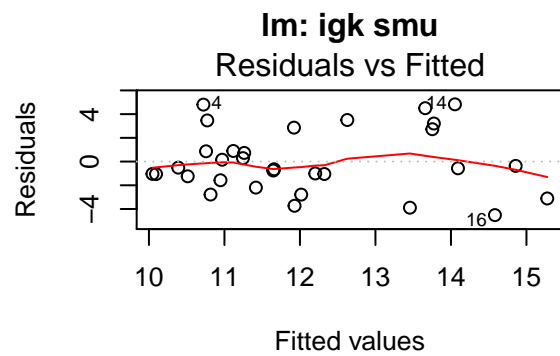
```
par(mfrow = c(2, 2))
lm = lm(miseq$iga_smu ~ Clin$`CD4 T cells (% viable)` + Clin$Age +
  Clin$Gender)
plot(lm, main = "lm: iga smu")
```



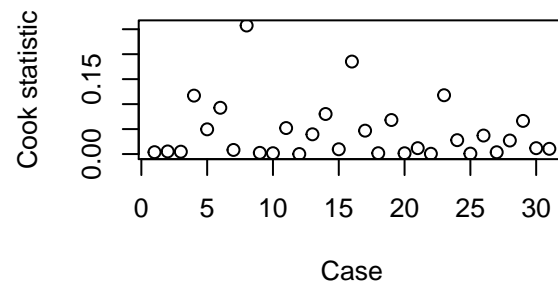
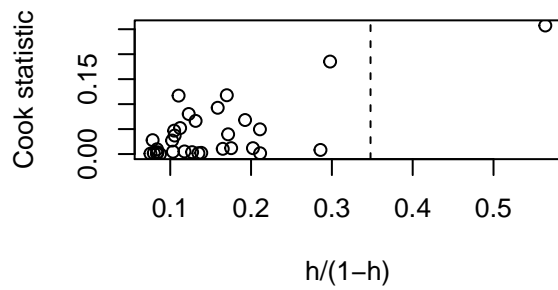
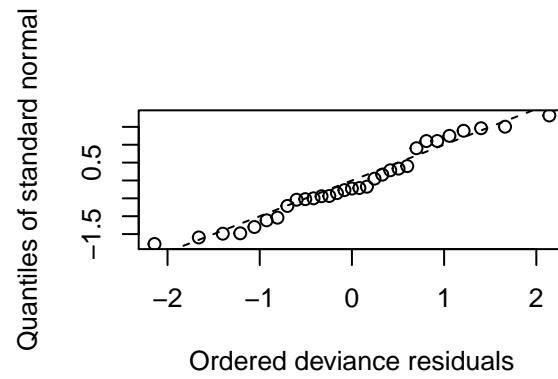
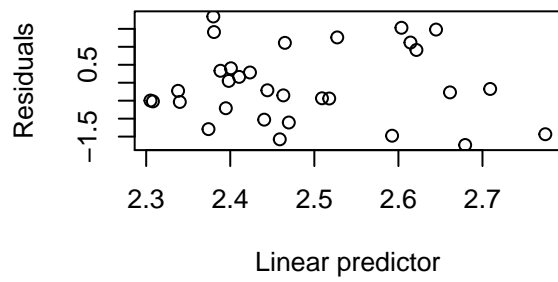
```
par(mfrow = c(2, 2))
glmGamma <- glm(miseq$iga_smu ~ Clin$`CD4 T cells (% viable)` +
  Clin$Age + Clin$Gender, family = Gamma(link = "log"))
glm.diag.plots(glmGamma)
```



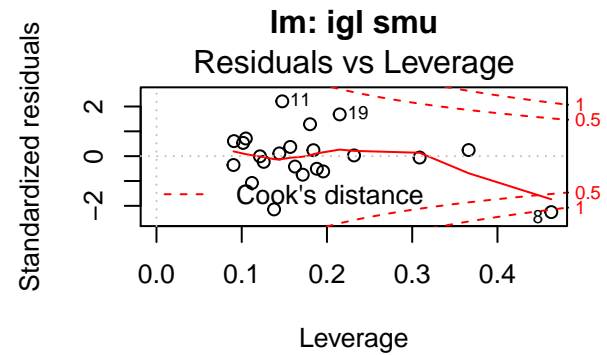
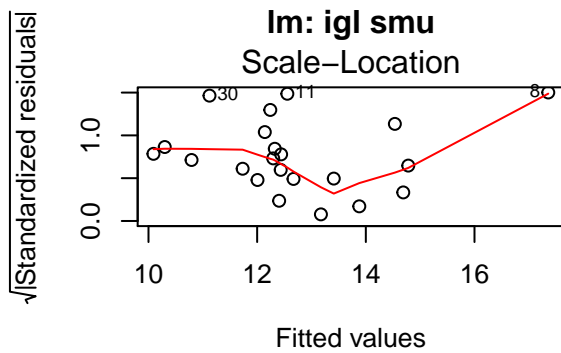
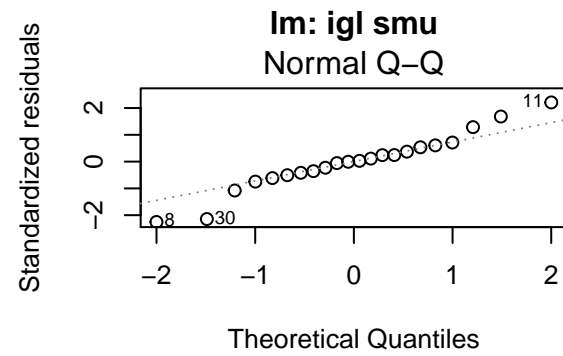
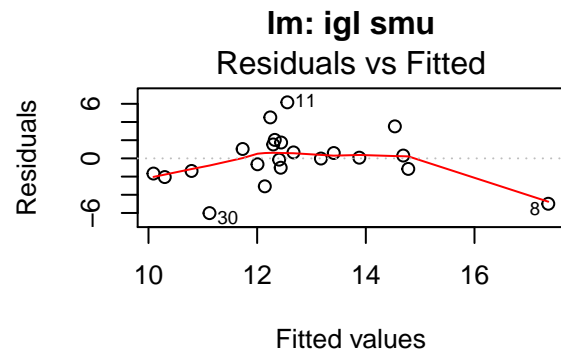
```
##### igk
par(mfrow = c(2, 2))
lm = lm(miseq$igk_smu ~ Clin$`CD4 T cells (% viable)` + Clin$Age +
  Clin$Gender)
plot(lm, main = "lm: igk smu")
```



```
par(mfrow = c(2, 2))
glmGamma <- glm(miseq$igk_smu ~ Clin$`CD4 T cells (% viable)` +
  Clin$Age + Clin$Gender, family = Gamma(link = "log"))
glm.diag.plots(glmGamma)
```



```
##### igl
par(mfrow = c(2, 2))
lm = lm(miseq$igl_smu ~ Clin$`CD4 T cells (% viable)` + Clin$Age +
        Clin$Gender)
plot(lm, main = "lm: igl smu")
```



```
par(mfrow = c(2, 2))
glmGamma <- glm(miseq$igl_smu ~ Clin$`CD4 T cells (% viable)` +
  Clin$Age + Clin$Gender, family = Gamma(link = "log"))
glm.diag.plots(glmGamma)
```

