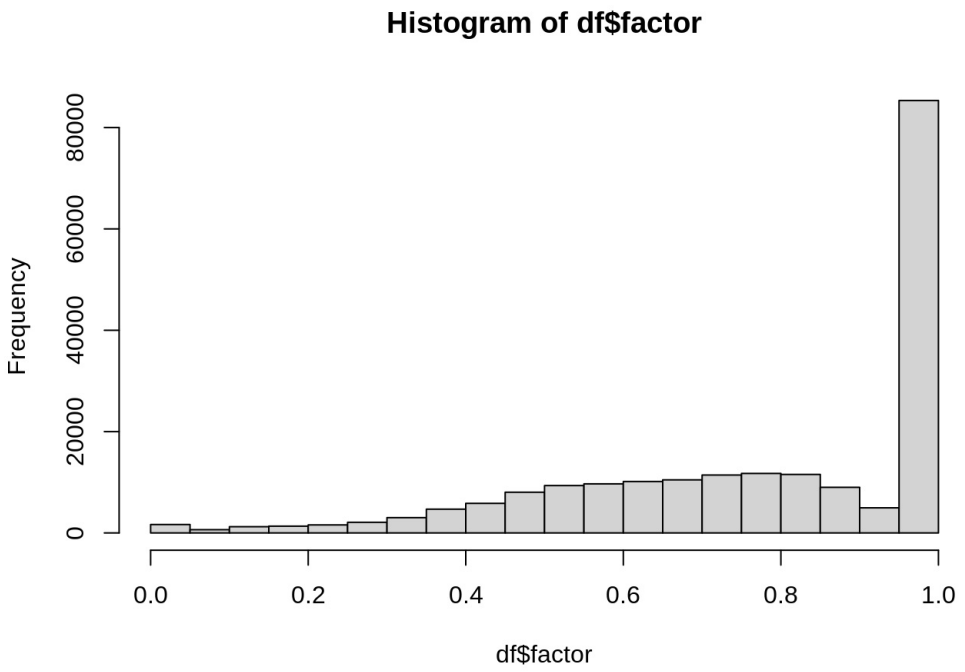
R Notebook

```
knitr::opts_chunk$set(echo = TRUE)
knitr::opts_knit$set(root.dir = '/fs/cbcb-lab/rob/students/noor/shoal_proj/')
knitr::opts_chunk$set(warning = FALSE, message = FALSE)
```

```
suppressPackageStartupMessages(source("run_limma.R"))
dirSal <- "/fs/cbcb-lab/rob/students/noor/shoal_proj/swim/sal_best/vbprior=1e0"
load("/fs/cbcb-lab/rob/students/noor/Uncertainity/mikelove-swimdown-216a1dd/simulate/data/simulate.rda")
```

Plotting factors

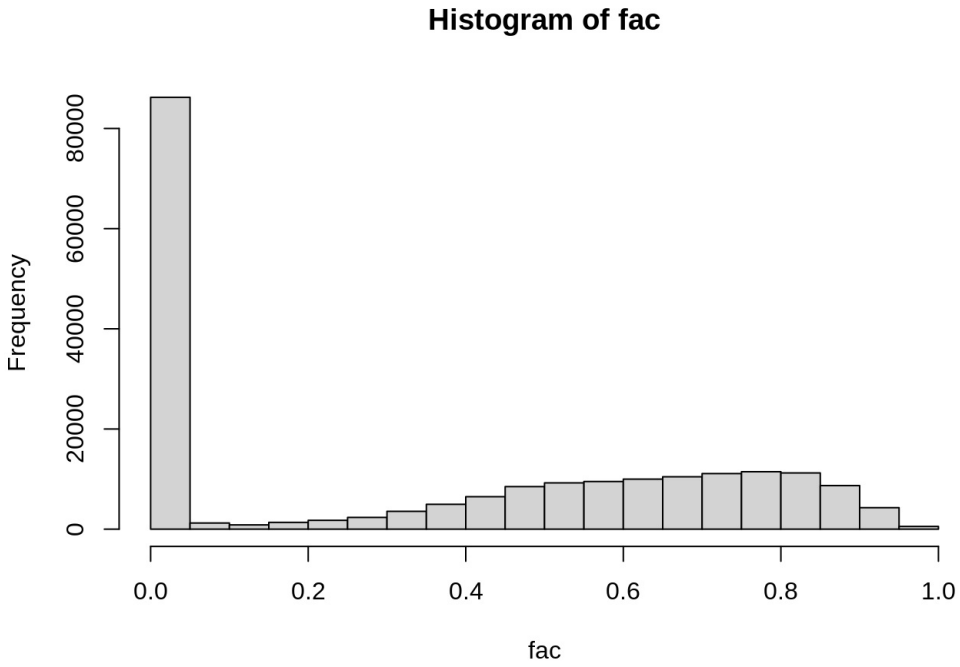
```
df <- read.delim("swim/shoal_best_updated/vbprior=1e0/c=0.1/prior.tsv")
hist(df$factor)
```



We have factors larger than 0.4 in great proportion

Normalizing data

```
salFiles <- getQuantFiles(dirSal, type="salmon")
dfSal <- tximport(salFiles[["files"]], txOut=T, type="salmon")
dg <- DGEList(dfSal$counts, group=salFiles$condition)
dg <- calcNormFactors(dg)
norm <- cpm(dg, log=F, normalized.lib.sizes=T)
fac <- computeFactor(norm)
hist(fac)
```



Plotting factors for equivalence classes

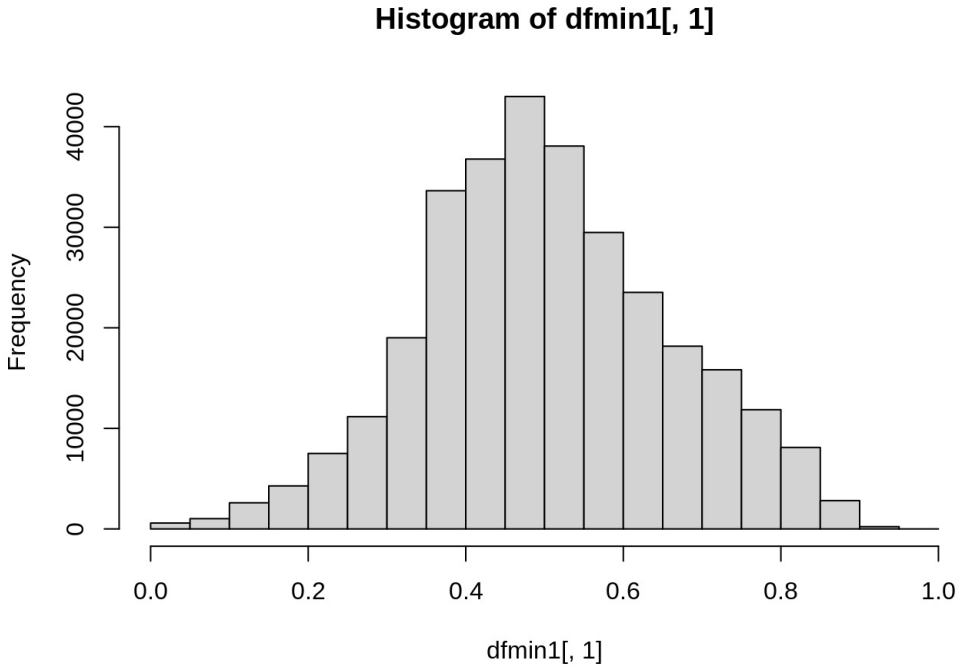
Minimum

```
dir <- "swim/shoal_best_updated/vbprior=1e0/c=100"

min1 <- file.path(dir, "fac_min.tsv")
min0 <- file.path(dir, "fac_min_updated.tsv")

dfmin1 <- read.delim(min1, header=F)
dfmin0 <- read.delim(min0, header=F)

hist(dfmin1[,1])
```




```
print(summary(dfmin1))
```

```
##        V1
##  Min.   :0.03628
##  1st Qu.:0.39632
##  Median :0.49275
##  Mean   :0.50277
##  3rd Qu.:0.60663
##  Max.   :0.95225
```

```
print(summary(dfmin0))
```

```
##        V1
##  Min.   :0.0000
##  1st Qu.:0.3960
##  Median :0.4927
##  Mean   :0.5023
##  3rd Qu.:0.6065
##  Max.   :0.9522
```

```
print(sum(abs(dfmin0[,1] - dfmin1[,1]) > 1e-3))
```

```
## [1] 302
```

```
print(sort(abs(dfmin0[,1] - dfmin1[,1]), decreasing = T)[1:20])
```

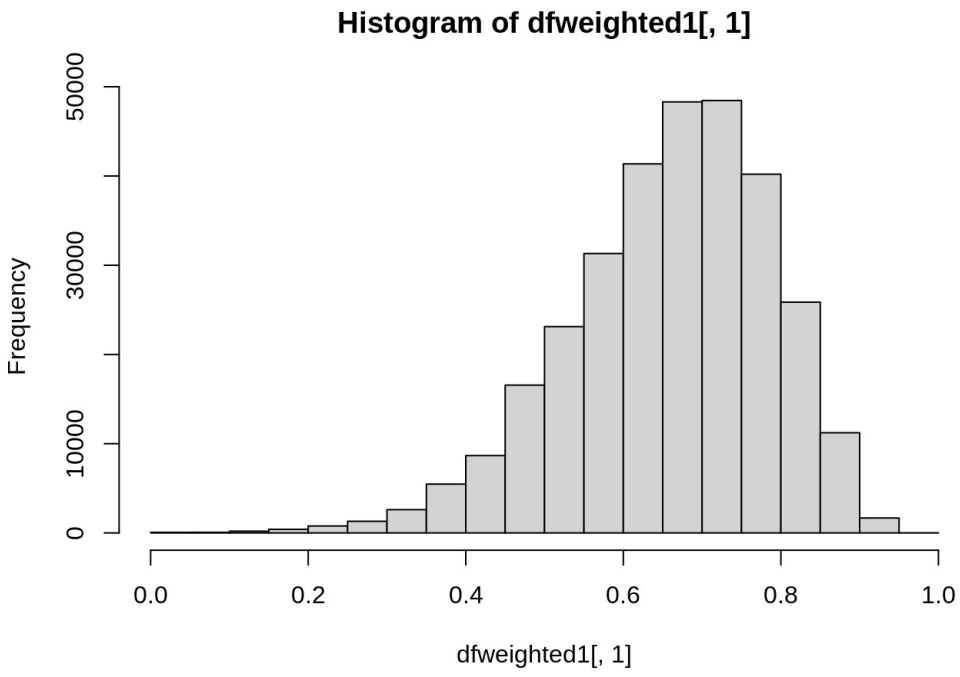
```
##  [1] 0.8928047 0.8671625 0.8566414 0.8315799 0.8267949 0.8075869 0.8054324
##  [8] 0.8034211 0.8030322 0.7998927 0.7991848 0.7949767 0.7949767 0.7949219
## [15] 0.7874311 0.7813473 0.7794500 0.7740111 0.7703706 0.7657056
```

Weighted

```
weighted1 <- file.path(dir, "fac_weighted.tsv")
weighted0 <- file.path(dir, "fac_weighted_updated.tsv")

dfweighted1 <- read.delim(weighted1, header=F)
dfweighted0 <- read.delim(weighted0, header=F)

hist(dfweighted1[,1])
```




```
print(summary(dfweighted1))
```

```
##        V1
##  Min.   :0.03628
##  1st Qu.:0.57986
##  Median :0.67325
##  Mean   :0.65882
##  3rd Qu.:0.75228
##  Max.   :0.97629
```

```
print(summary(dfweighted0))
```

```
##        V1
##  Min.   :0.03628
##  1st Qu.:0.57984
##  Median :0.67324
##  Mean   :0.65879
##  3rd Qu.:0.75225
##  Max.   :0.97629
```

```
print(sum(abs(dfweighted0[,1] - dfweighted1[,1]) > 1e-3))
```

```
## [1] 126
```

```
print(sort(abs(dfweighted0[,1] - dfweighted1[,1]), decreasing = T)[1:20])
```

```
##  [1] 0.8901140 0.7570520 0.5597600 0.5152360 0.4353930 0.4335650 0.4118220
##  [8] 0.4044570 0.3407800 0.2457368 0.2299520 0.2097730 0.2076100 0.1865530
## [15] 0.1745620 0.1733250 0.1586070 0.1336170 0.0807453 0.0783299
```