

How to calculate the mean

KE Lotterhos

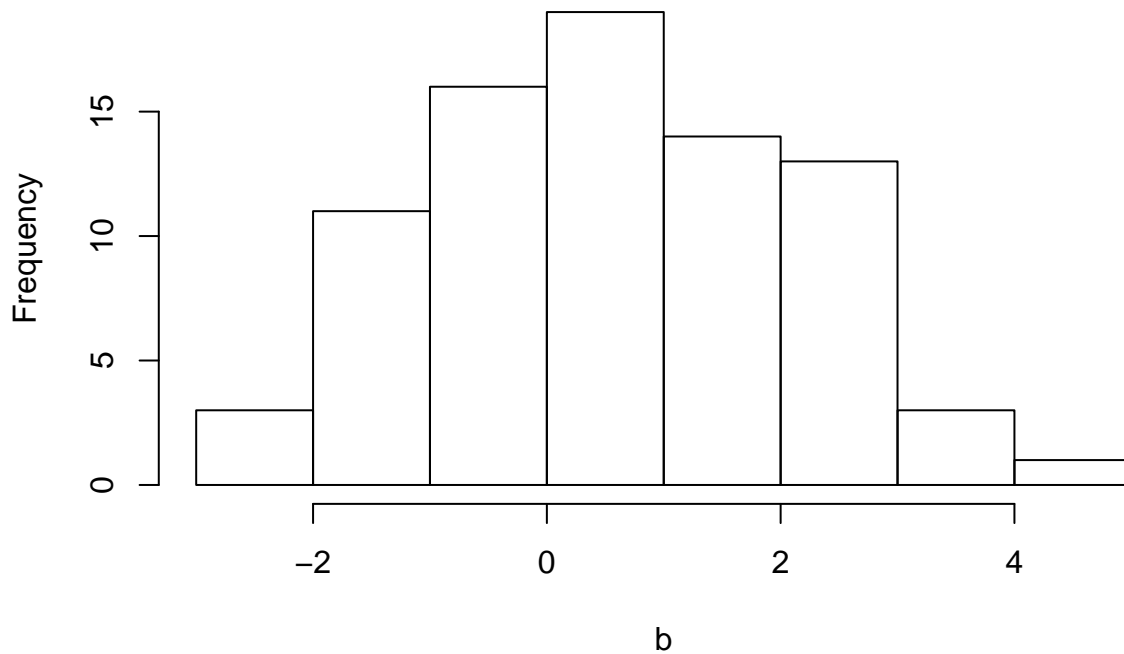
3/9/2021

```
# Different ways of calculating means ####
getmeans <- function(ind=1:length(a), true){
  obs_all <- mean(b[ind]) # This may or may not be what your eMM does
  obs_mean_GE <- mean(tapply(b[ind],a[ind], mean))
  obs_mean_G <- mean(tapply(b[ind],G_lev[ind], mean))
  obs_mean_E <- mean(tapply(b[ind],E_lev[ind], mean))
  obs_mean_cGE <- mean(c(obs_mean_G, obs_mean_E)) # This may or may not be what your code does.
  out <- cbind(obs_all, obs_mean_GE, obs_mean_G,obs_mean_E, obs_mean_cGE)
  round(out - true, 4)
}
```

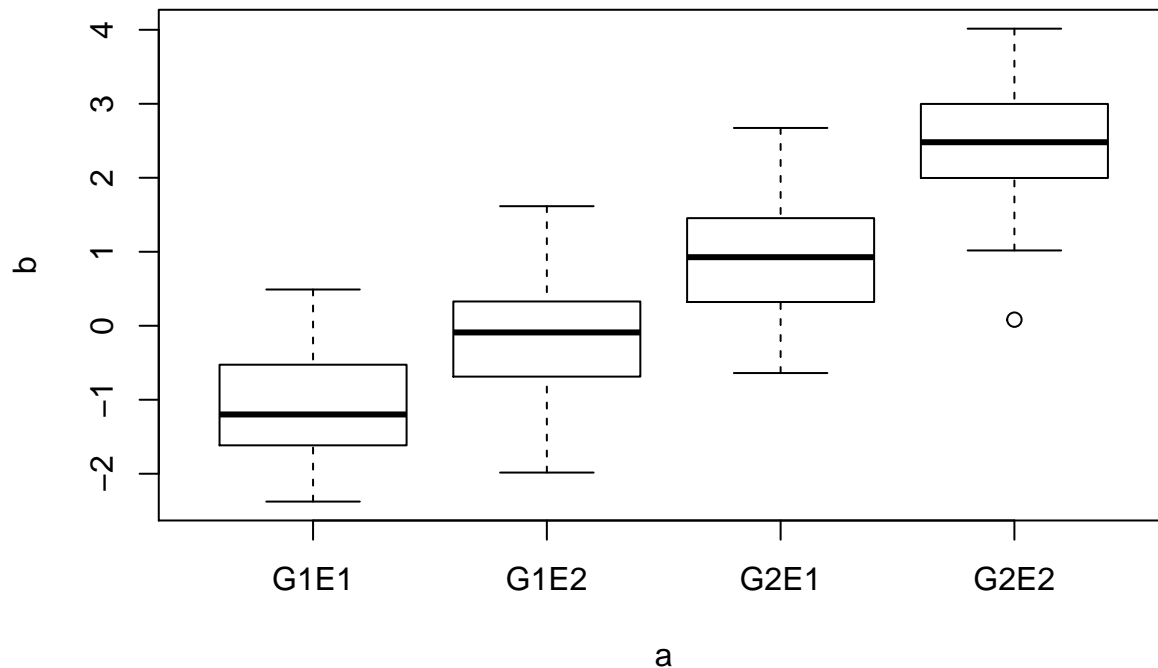
consider 2x2 case

```
a <- rep(c("G1E1","G1E2", "G2E1", "G2E2"), each=20)
G_lev <- substr(a, start=1, stop=2)
E_lev <- substr(a, start=3, stop=4)
b <- c(rnorm(length(a)/4, -1), rnorm(length(a)/4, 0), rnorm(length(a)/4, 1), rnorm(length(a)/4, 2))
hist(b)
```

Histogram of b



```
boxplot(b~a)
```



```

true1 <- mean(c(-1,0,1,2)) # true mean

getmeans(1:length(a), true1) # no imbalance

##      obs_all obs_mean_GE obs_mean_G obs_mean_E obs_mean_cGE
## [1,]  0.0498    0.0498    0.0498    0.0498    0.0498

getmeans(15:length(a), true1) # imbalance in first category

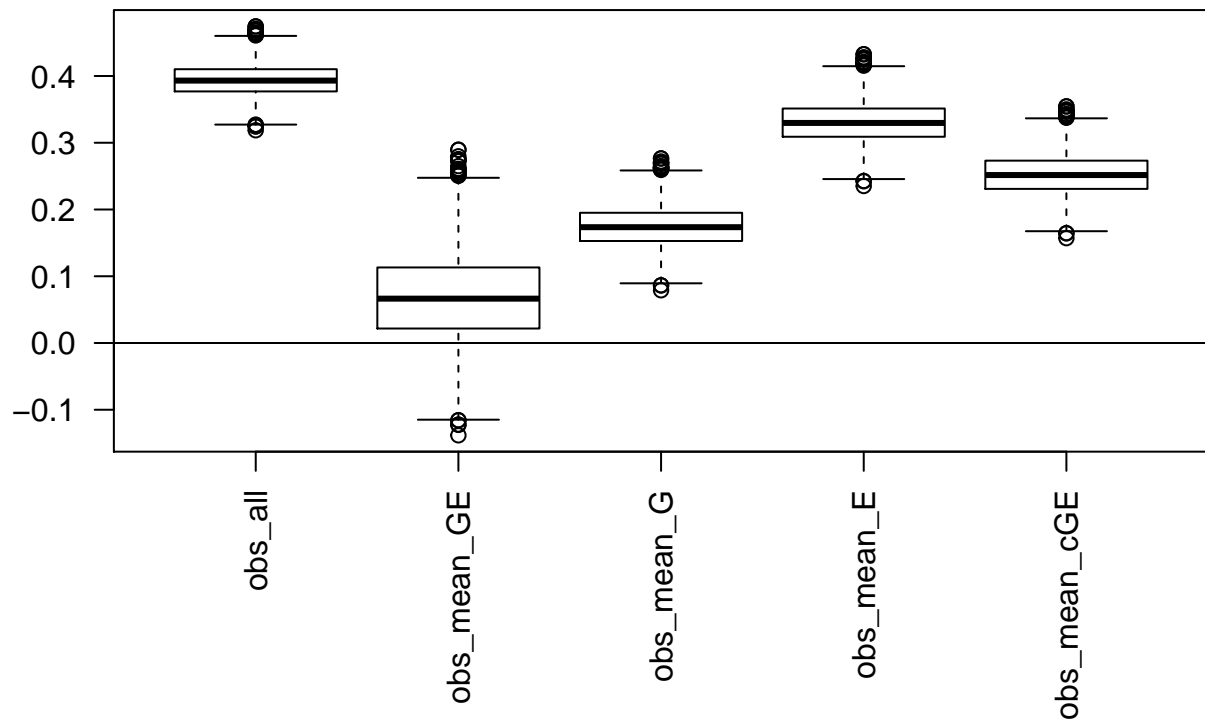
##      obs_all obs_mean_GE obs_mean_G obs_mean_E obs_mean_cGE
## [1,]  0.4072    0.1051    0.1913    0.3475    0.2694

getmeans(1:(length(a)-15), true1) # imbalance in last category

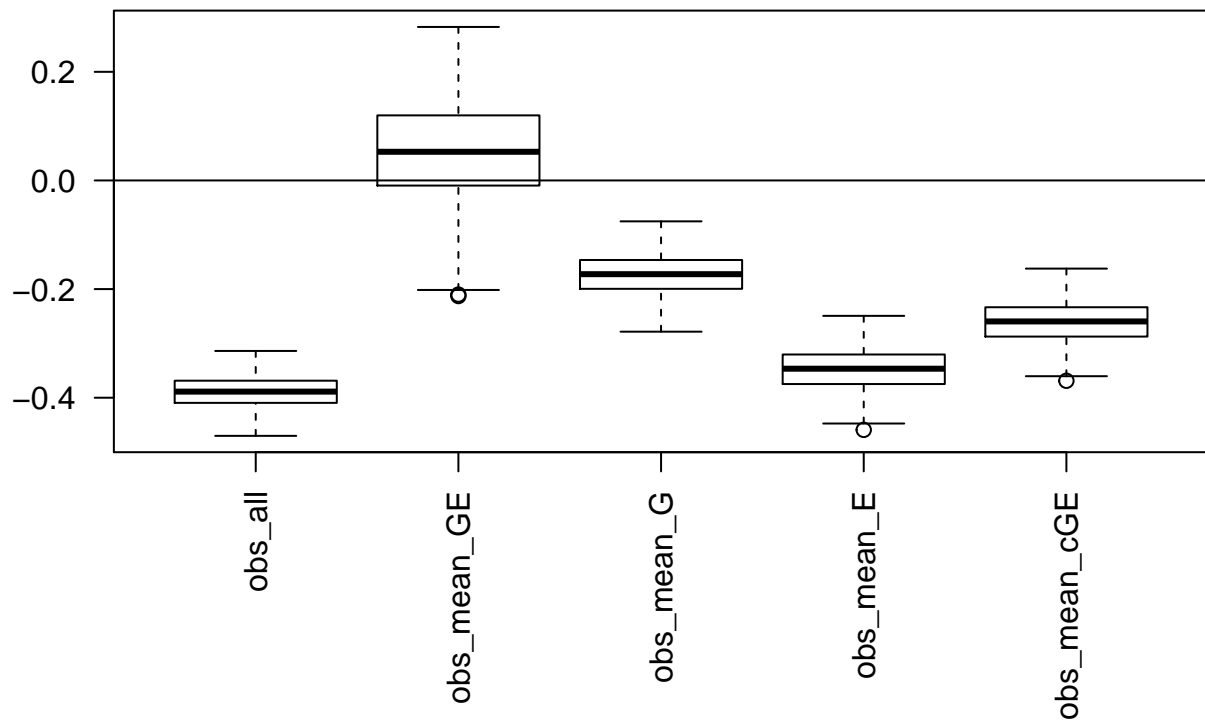
##      obs_all obs_mean_GE obs_mean_G obs_mean_E obs_mean_cGE
## [1,] -0.3766    0.0788   -0.1568   -0.3308   -0.2438

par(mar=c(10,3,1,1))
bob <- replicate(10000, getmeans(c(sample(1:20,5), 20:length(a)), true1), simplify = TRUE)
rownames(bob) <- c("obs_all", "obs_mean_GE", "obs_mean_G", "obs_mean_E", "obs_mean_cGE")
#head(bob)
boxplot(t(bob), las=2)
abline(h=0)

```



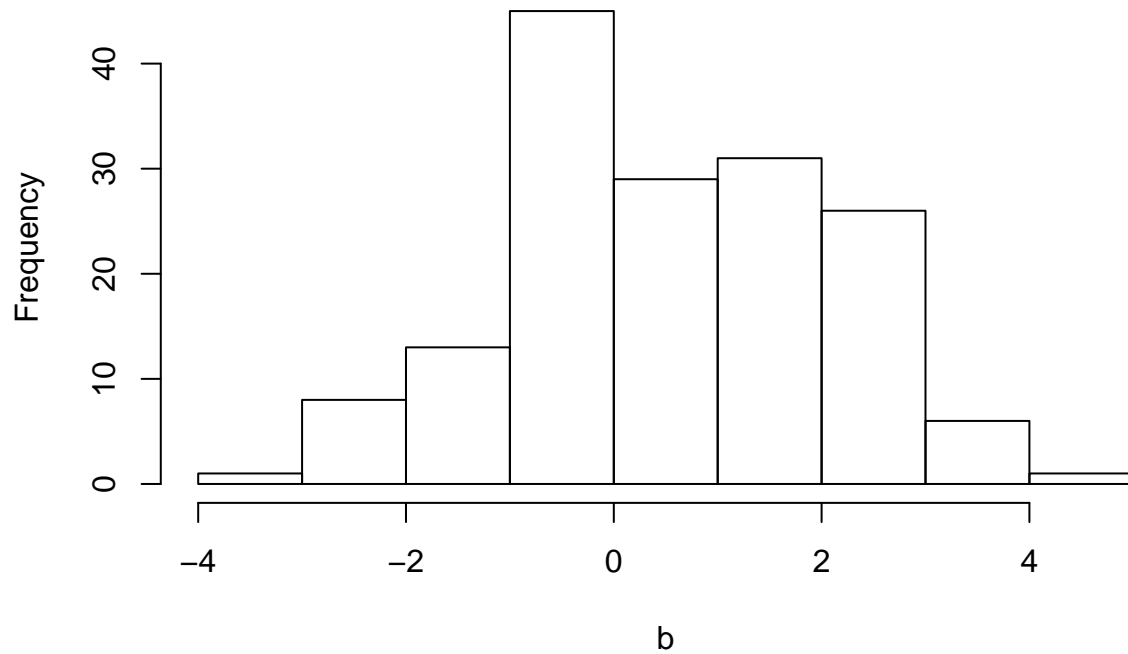
```
bob <- replicate(1000, getmeans(c(1:(length(a)-20), sample((length(a)-20):length(a), 5)), true1), simplify)
#head(bob)
rownames(bob) <- c("obs_all", "obs_mean_GE", "obs_mean_G", "obs_mean_E", "obs_mean_cGE")
#head(bob)
boxplot(t(bob), las=2)
abline(h=0)
```



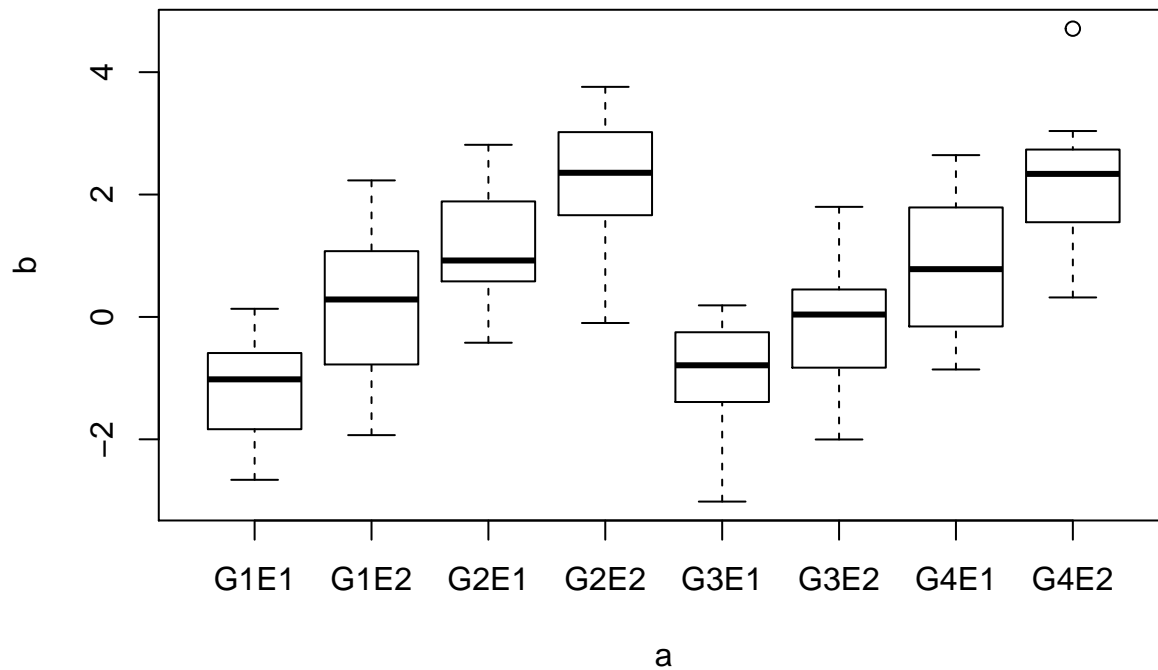
consider 2x4 common garden case

```
a <- rep(c("G1E1","G1E2", "G2E1", "G2E2", "G3E1", "G3E2", "G4E1", "G4E2"), each=20)
G_lev <- substr(a, start=1, stop=2)
E_lev <- substr(a, start=3, stop=4)
b <- c(rnorm(length(a)/8, -1),
      rnorm(length(a)/8, 0),
      rnorm(length(a)/8, 1),
      rnorm(length(a)/8, 2),
      rnorm(length(a)/8, -1),
      rnorm(length(a)/8, 0),
      rnorm(length(a)/8, 1),
      rnorm(length(a)/8, 2))
hist(b)
```

Histogram of b



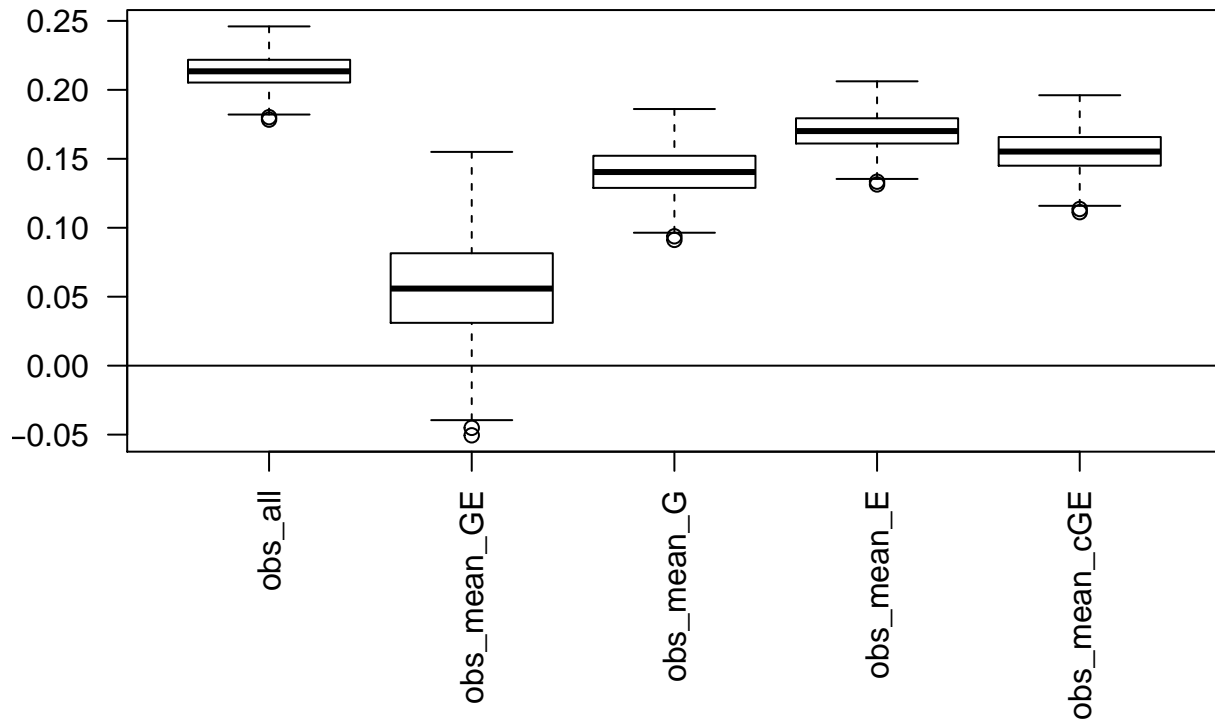
```
boxplot(b~a)
```



```
par(mar=c(10,3,1,1))

true2 <- mean(c(-1,0,1,2, -1, 0, 1,2,-1,0,1,2)) # true mean

bob <- replicate(1000,getmeans(c(sample(1:20,5),20:length(a)), true2),
                    simplify = TRUE)
rownames(bob) <- c("obs_all", "obs_mean_GE", "obs_mean_G", "obs_mean_E", "obs_mean_cGE")
#head(bob)
boxplot(t(bob), las=2)
abline(h=0)
```



```

bob <- replicate(1000, getmeans(c(1:(length(a)-20),
                                sample((length(a)-20):length(a), 5)), true2),
                simplify = TRUE)
rownames(bob) <- c("obs_all", "obs_mean_GE", "obs_mean_G", "obs_mean_E", "obs_mean_cGE")
#head(bob)
boxplot(t(bob), las=2)
abline(h=0)

```

