

Simulation Study

Michael Sachs

2015-07-29

Consistency of Maximum Likelihood Estimates

```
library(parallel)
library(pcrmeta)
nnns <- c(10, 23, 50, 100)
truth <- c(.38, -.18, .0452, .1119, .0075)

run_bias <- function(n){

  dat <- gen_dataset(n)
  fit <- tryCatch(fit_model4(dat$Yi, dat$Xi, dat$sigma.hr, dat$sigma.pcr), error = function(e) list(par
  fit$par

}

biasres <- vector("list", 4)

for(i in 1:4){
  biasres[[i]] <- matrix(unlist(mclapply(1:2000, function(j) run_bias(nnns[i]))), ncol = 5, byrow = TRUE
  colnames(biasres[[i]]) <- c("mux", "muy", "beta", "vmx", "vmg")
}
```

All distributions should be centered around 0. The red line in the histograms is the true value.

```
for(i in 1:4){

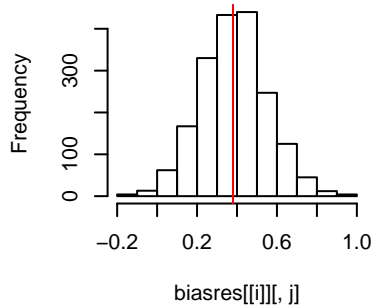
  print(paste("N =", nnns[i]))
  print(summary(biasres[[i]]))
  par(mfrow = c(2, 3))
  for(j in 1:5){
    hist(biasres[[i]][, j], main = paste0(colnames(biasres[[i]])[j], "; N = ", nnns[i]))
    abline(v = truth[j], col = "red")
  }

}
```

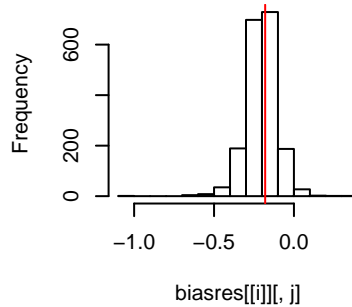
```
## [1] "N = 10"
##      mux      muy      beta      vmx
## Min.   :-0.1535 Min.   :-1.0653 Min.   :-1.37632 Min.   :0.0000
## 1st Qu.: 0.2732 1st Qu.: -0.2529 1st Qu.: -0.04441 1st Qu.: 0.1279
## Median : 0.3850 Median : -0.1996 Median : 0.04915 Median : 0.2025
## Mean   : 0.3847 Mean   : -0.2025 Mean   : 0.05372 Mean   : 0.2217
## 3rd Qu.: 0.4891 3rd Qu.: -0.1439 3rd Qu.: 0.14341 3rd Qu.: 0.2927
## Max.   : 0.9798 Max.   : 0.3947 Max.   : 1.31496 Max.   : 1.0194
```

```
## NA's :118      NA's :118      NA's :118      NA's :118
##      vmg
## Min. :0.00000
## 1st Qu.:0.00000
## Median :0.00795
## Mean :0.01254
## 3rd Qu.:0.01980
## Max. :0.09397
## NA's :118
```

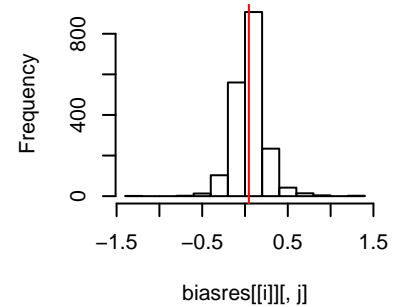
mux; N = 10



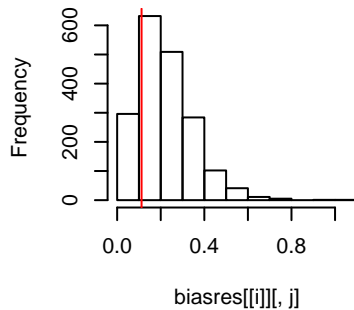
muy; N = 10



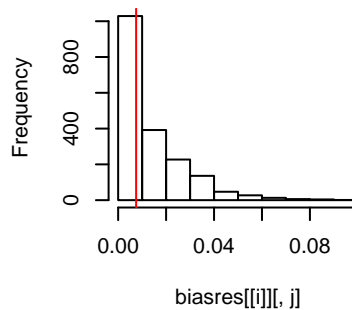
beta; N = 10



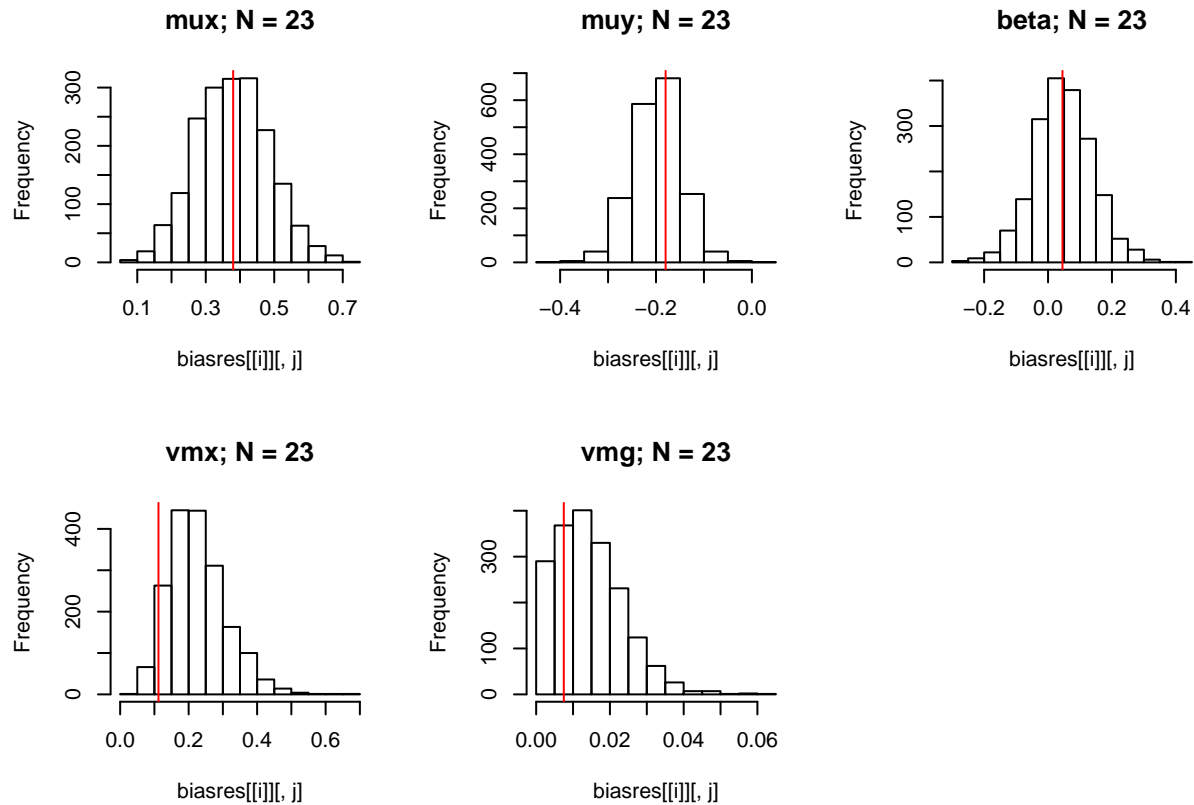
vmx; N = 10



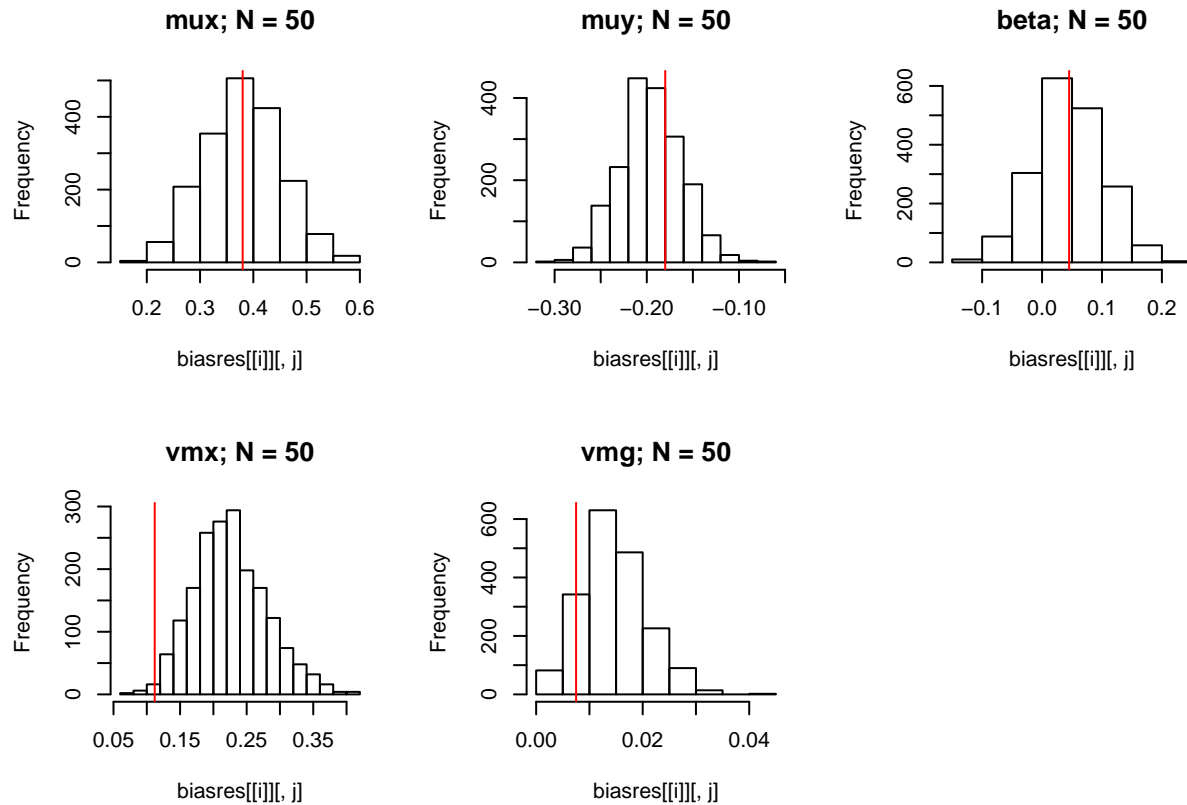
vmg; N = 10



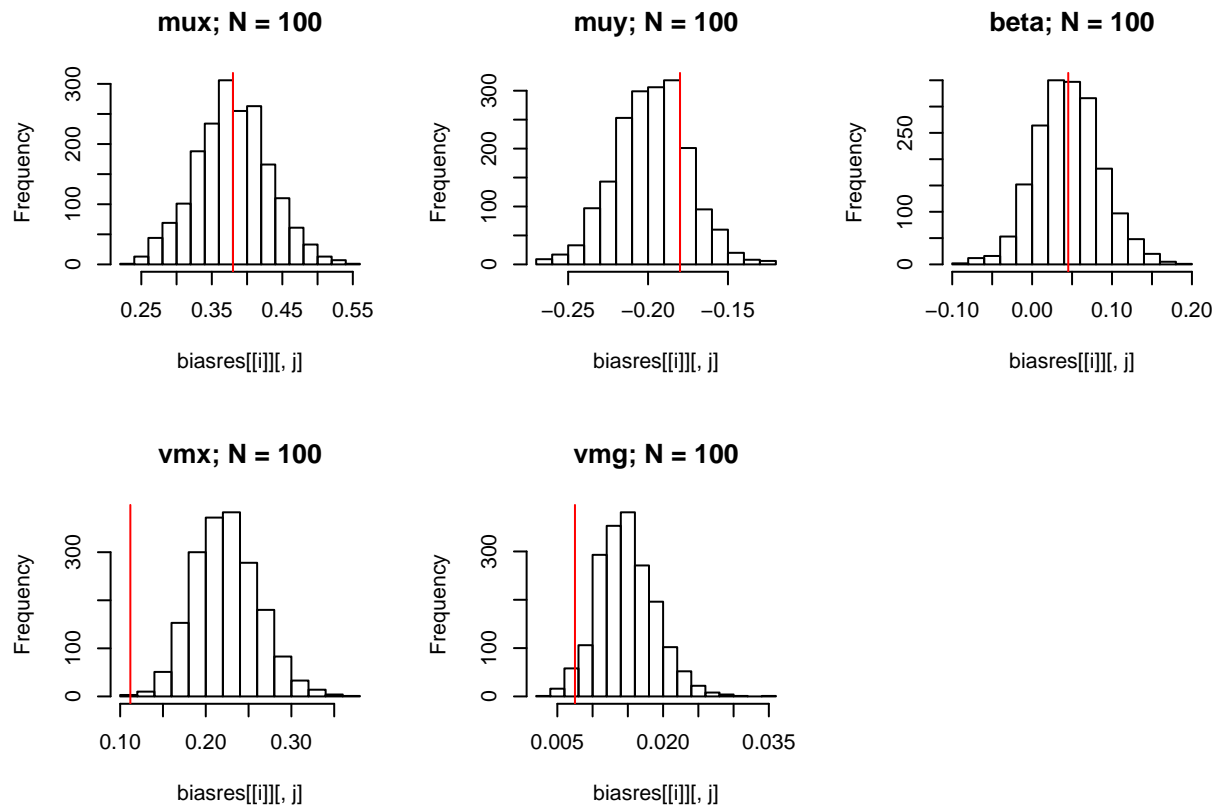
```
## [1] "N = 23"
##      mux      muy      beta      vmx
## Min. :0.08059 Min. : -0.4242 Min. : -0.29898 Min. :0.04614
## 1st Qu.:0.30330 1st Qu.: -0.2305 1st Qu.: -0.01243 1st Qu.:0.16473
## Median :0.37606 Median : -0.1974 Median : 0.04526 Median :0.21660
## Mean :0.37830 Mean : -0.1985 Mean : 0.04682 Mean :0.22548
## 3rd Qu.:0.45067 3rd Qu.: -0.1653 3rd Qu.: 0.10589 3rd Qu.:0.27390
## Max. :0.73902 Max. : 0.0052 Max. : 0.41855 Max. :0.65944
## NA's :150      NA's :150      NA's :150      NA's :150
##      vmg
## Min. :0.00000
## 1st Qu.:0.00768
## Median :0.01316
## Mean :0.01432
## 3rd Qu.:0.01992
## Max. :0.06269
## NA's :150
```



```
## [1] "N = 50"
##      mux      muy      beta      vmx
## Min.   :0.1904  Min.   : -0.30542  Min.   : -0.13137  Min.   :0.07597
## 1st Qu.:0.3265  1st Qu.: -0.21733  1st Qu.: 0.00793   1st Qu.:0.18788
## Median :0.3787  Median : -0.19593  Median : 0.04306   Median :0.22111
## Mean   :0.3810  Mean   : -0.19538  Mean   : 0.04474   Mean   :0.22526
## 3rd Qu.:0.4326  3rd Qu.: -0.17381  3rd Qu.: 0.08453   3rd Qu.:0.26005
## Max.   :0.5927  Max.   : -0.07918  Max.   : 0.22810   Max.   :0.40279
## NA's   :128    NA's   :128      NA's   :128      NA's   :128
##      vmg
## Min.   :0.00000
## 1st Qu.:0.01036
## Median :0.01383
## Mean   :0.01448
## 3rd Qu.:0.01803
## Max.   :0.04169
## NA's   :128
```



```
## [1] "N = 100"
##      mux      muy      beta      vmx
## Min.   :0.2336  Min.   :-0.2690  Min.   :-0.08459  Min.   :0.1114
## 1st Qu.:0.3453  1st Qu.: -0.2125  1st Qu.: 0.01771  1st Qu.:0.1974
## Median :0.3783  Median : -0.1974  Median : 0.04401  Median :0.2221
## Mean   :0.3798  Mean   : -0.1976  Mean   : 0.04557  Mean   :0.2237
## 3rd Qu.:0.4147  3rd Qu.: -0.1823  3rd Qu.: 0.07273  3rd Qu.:0.2473
## Max.   :0.5414  Max.   : -0.1205  Max.   : 0.18069  Max.   :0.3621
## NA's   :135    NA's   :135    NA's   :135    NA's   :135
##      vmg
## Min.   :0.00275
## 1st Qu.:0.01197
## Median :0.01449
## Mean   :0.01477
## 3rd Qu.:0.01721
## Max.   :0.03504
## NA's   :135
```



Confidence and Prediction Interval Coverage

```
splank <- function(fun){
  matrix(unlist(mclapply(1:2000, fun)), ncol = 2, byrow = TRUE)
}

tryrun <- function(i, n, ...){
  tryCatch(run_pred_coverage(i = i, n = n, ...), error = function(e) c(NA, NA))
}

covers <- list(cover10 = splank(function(i) tryrun(i, n = 10)),
  cover23 = splank(function(i) tryrun(i, n = 23)),

  cover10t = splank(function(i) tryrun(i, n = 10, zst = -qt(0.025, df = 10 - 1))),
  cover23t = splank(function(i) tryrun(i, n = 23, zst = -qt(0.025, df = 23 - 1))))

lapply(covers, colMeans, na.rm = TRUE)

## $cover10
## [1] 0.7735546 0.8847185
##
```

```
## $cover23
## [1] 0.8731020 0.9045553
##
## $cover10t
## [1] 0.8172157 0.9239766
##
## $cover23t
## [1] 0.8703404 0.9108590
```