

negativebinomial_qc_vs_lme4_glmer

sarah

10/26/2021

I simulated a single dataset with sample size 10,000 and cluster size 5 under our negative binomial random intercept GLMM with $\beta \sim \text{Uniform}(-0.2, 0.2)$, single variance component $\text{Sigma} = 0.01$ and $r = 10$. Here I compare the fit of our NB regression with two most popular R packages for fitting Negative Binomial GLMM's with options to estimate "r":

1. lme4: The function `glmer.nb` estimates r pretty well, but very slowly. To get confidence intervals, it takes a significant amount of additional time, but no inference on r .
2. glmmTMB: The function `glmmTMB` estimates r pretty well and a little faster than lme4 but still slow. We can get the CI for all parameters in much less time than lme4.

lme4: glmer.nb

```
# Start the clock!
ptm <- proc.time()
m.nb <- glmer.nb(Y ~ 1 + X2 + X3 + (1|group), data = data, verbose = TRUE)

## theta.ml: iter 0 'theta = 1.029180'

## theta.ml: iter1 theta =1.53524

## theta.ml: iter2 theta =2.22902

## theta.ml: iter3 theta =3.16185

## theta.ml: iter4 theta =4.39397

## theta.ml: iter5 theta =5.9918

## theta.ml: iter6 theta =8.01926

## theta.ml: iter7 theta =10.5174

## theta.ml: iter8 theta =13.4625

## theta.ml: iter9 theta =16.6897

## theta.ml: iter10 theta =19.7865

## theta.ml: iter11 theta =22.0722

## theta.ml: iter12 theta =23.0503

## theta.ml: iter13 theta =23.1898

## theta.ml: iter14 theta =23.1922

## theta.ml: iter15 theta =23.1922

## th := est_theta(glmer(..)) = 23.19225 --> dev.= -2*logLik(.) = 135769.6
## 1: th= 11.42283134, dev=135721.37334302, beta[1]= 0.03555525
## 2: th= 47.08817777, dev=135825.00632533, beta[1]= 0.02923654
## 3: th= 4.759896381, dev=135906.62093146, beta[1]= 0.04567041
## 4: th= 17.27769222, dev=135744.76706271, beta[1]= 0.03261097
## 5: th= 8.176347719, dev=135732.60433672, beta[1]= 0.03925245
## 6: th= 11.11006939, dev=135720.88078436, beta[1]= 0.03580499
## 7: th= 10.68397568, dev=135720.53531345, beta[1]= 0.03617708
## 8: th= 10.53962030, dev=135720.51700797, beta[1]= 0.03630743
## 9: th= 10.56141548, dev=135720.51624516, beta[1]= 0.03629208
## 10: th= 10.56362598, dev=135720.51624051, beta[1]= 0.03628591
## 11: th= 10.56329011, dev=135720.51624070, beta[1]= 0.03628591
## 12: th= 10.56381741, dev=135720.51624079, beta[1]= 0.03628591
## 13: th= 10.56362598, dev=135720.51624065, beta[1]= 0.03628591
```

```
# Stop the clock
proc.time() - ptm
```

```
##      user  system elapsed
## 109.199    2.646 112.850
```

Show estimates

```
m.nb
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: Negative Binomial(10.5636) ( log )
## Formula: Y ~ 1 + X2 + X3 + (1 | group)
## Data: data
##      AIC      BIC    logLik deviance df.resid
## 135730.52 135774.62 -67860.26 135720.52    49995
## Random effects:
## Groups Name      Std.Dev.
## group (Intercept) 0.1318
## Number of obs: 50000, groups: group, 10000
## Fixed Effects:
## (Intercept)          X2          X3
##      0.03629      0.10488      0.02267
```

Show estimated r from lme4: glmer.nb

```
getME(m.nb, "glmer.nb.theta")
```

```
## [1] 10.56363
```

Show confidence intervals from lme4: glmer.nb

```
ptm <- proc.time()
confint(m.nb)
```

```
## Computing profile confidence intervals ...
```

```
##           2.5 %      97.5 %
## .sig01      0.10538029 0.15424433
## (Intercept) 0.02629848 0.04619829
## X2          0.09588142 0.11388454
## X3          0.01366931 0.03167231
```

```
# Stop the clock
proc.time() - ptm
```

```
##      user  system elapsed
## 236.851    6.126 243.086
```

glmmTMB: glmmTMB

```
# Start the clock!
ptm <- proc.time()
m.glmmtmb_nb <- glmmTMB(Y ~ 1 + X2 + X3 + (1|group), data = data, family=nbinom2)
# Stop the clock
proc.time() - ptm
```

```
##      user  system elapsed
## 115.330   0.985 116.421
```

Show estimates

```
m.glmmtmb_nb

## Formula:          Y ~ 1 + X2 + X3 + (1 | group)
## Data: data
##      AIC      BIC    logLik df.resid
## 135731.8 135775.9 -67860.9   49995
## Random-effects (co)variances:
##
## Conditional model:
##  Groups Name      Std.Dev.
##  group (Intercept) 0.1288
##
## Number of obs: 50000 / Conditional model: group, 10000
##
## Dispersion parameter for nbinom2 family (): 10.5
##
## Fixed Effects:
##
## Conditional model:
## (Intercept)          X2          X3
##    0.03739      0.10495    0.02269
```

Show estimated r from glmmTMB

```
sigma(m.glmmtmb_nb)
```

```
## [1] 10.50208
```

Show confidence intervals from glmmTMB

```
ptm <- proc.time()
confint(m.glmmtmb_nb, full = TRUE)
```

```
##              2.5 %      97.5 %    Estimate
## cond.(Intercept) 0.02740384 0.04736648 0.03738516
## cond.X2          0.09594516 0.11396380 0.10495448
## cond.X3          0.01367701 0.03169594 0.02268648
## sigma           8.94628384 12.32842980 10.50207752
## group.cond.Std.Dev.(Intercept) 0.10424123 0.15906907 0.12876939
```

```
# Stop the clock  
proc.time() - ptm
```

```
##      user  system elapsed  
##    0.673    0.031    0.704
```