Analysis of BMI GIANT GWAS data

Simina Boca, Jeff Leek

2016-10-27

Load the relevant libraries:

```
library(readr)
library(dplyr)
library(fdrtool)
library(betareg)
library(splines)
library(Hmisc)
library(ggplot2)
library(reshape2)
library(swfdr)
```

Load the .RData file with the BMI GIANT GWAS meta-analysis data:

```
load("BMI_GIANT_GWAS.RData")

tot <- BMI_GIANT_GWAS
```

Estimate fraction of true null hypotheses in a regression framework

Create the design matrix, using natural cubic splines with 5 degrees of freedom to model `N` and 3 discrete categories for the MAFs:

```
X <- model.matrix(~ splines::ns(N,5) + Freq_MAF_Int_Hapmap, data = tot)[,-1]
dim(X)
```

```
## [1] 2500573       7
```

```
head(X)
```

```
##   splines::ns(N, 5)1 splines::ns(N, 5)2 splines::ns(N, 5)3
## 1       4.414107e-01       5.538398e-01      -0.0017421409
## 2       0.000000e+00       3.954615e-10      -0.1655612193
## 3       3.884106e-05       9.880678e-01       0.0099138318
## 4       3.209714e-04       9.891039e-01       0.0088153851
## 5       9.327150e-02       9.061998e-01       0.0002901038
## 6       6.724476e-04       9.894590e-01       0.0082264435
##   splines::ns(N, 5)4 splines::ns(N, 5)5 Freq_MAF_Int_Hapmap[0.127,0.302)
## 1       0.0034871548      -0.0017450139                                 0
## 2       0.3336072837       0.8319539352                                 0
## 3       0.0039611703      -0.0019816905                                 1
## 4       0.0035221381      -0.0017623939                                 1
## 5       0.0002264943      -0.0001133405                                 1
## 6       0.0032867973      -0.0016447160                                 0
```

```
##   Freq_MAF_Int_Hapmap[0.302,0.500]
## 1                                1
## 2                                1
## 3                                0
## 4                                0
## 5                                0
## 6                                1
```

Run code to estimate the fraction of true null hypotheses within a regression framework with the design matrix specified above:

```
pi0Est <- lm_pi0(pValues=tot$p, X=X, smooth.df=3)
```

```
## [1] "At test #: 10000"
## [1] "At test #: 20000"
## [1] "At test #: 30000"
## [1] "At test #: 40000"
## [1] "At test #: 50000"
## [1] "At test #: 60000"
## [1] "At test #: 70000"
## [1] "At test #: 80000"
## [1] "At test #: 90000"
## [1] "At test #: 100000"
## [1] "At test #: 110000"
## [1] "At test #: 120000"
## [1] "At test #: 130000"
## [1] "At test #: 140000"
## [1] "At test #: 150000"
## [1] "At test #: 160000"
## [1] "At test #: 170000"
## [1] "At test #: 180000"
## [1] "At test #: 190000"
## [1] "At test #: 200000"
## [1] "At test #: 210000"
## [1] "At test #: 220000"
## [1] "At test #: 230000"
## [1] "At test #: 240000"
## [1] "At test #: 250000"
## [1] "At test #: 260000"
## [1] "At test #: 270000"
## [1] "At test #: 280000"
## [1] "At test #: 290000"
## [1] "At test #: 300000"
## [1] "At test #: 310000"
## [1] "At test #: 320000"
## [1] "At test #: 330000"
## [1] "At test #: 340000"
## [1] "At test #: 350000"
## [1] "At test #: 360000"
## [1] "At test #: 370000"
## [1] "At test #: 380000"
## [1] "At test #: 390000"
## [1] "At test #: 400000"
```

```
## [1] "At test #: 410000"
## [1] "At test #: 420000"
## [1] "At test #: 430000"
## [1] "At test #: 440000"
## [1] "At test #: 450000"
## [1] "At test #: 460000"
## [1] "At test #: 470000"
## [1] "At test #: 480000"
## [1] "At test #: 490000"
## [1] "At test #: 500000"
## [1] "At test #: 510000"
## [1] "At test #: 520000"
## [1] "At test #: 530000"
## [1] "At test #: 540000"
## [1] "At test #: 550000"
## [1] "At test #: 560000"
## [1] "At test #: 570000"
## [1] "At test #: 580000"
## [1] "At test #: 590000"
## [1] "At test #: 600000"
## [1] "At test #: 610000"
## [1] "At test #: 620000"
## [1] "At test #: 630000"
## [1] "At test #: 640000"
## [1] "At test #: 650000"
## [1] "At test #: 660000"
## [1] "At test #: 670000"
## [1] "At test #: 680000"
## [1] "At test #: 690000"
## [1] "At test #: 700000"
## [1] "At test #: 710000"
## [1] "At test #: 720000"
## [1] "At test #: 730000"
## [1] "At test #: 740000"
## [1] "At test #: 750000"
## [1] "At test #: 760000"
## [1] "At test #: 770000"
## [1] "At test #: 780000"
## [1] "At test #: 790000"
## [1] "At test #: 800000"
## [1] "At test #: 810000"
## [1] "At test #: 820000"
## [1] "At test #: 830000"
## [1] "At test #: 840000"
## [1] "At test #: 850000"
## [1] "At test #: 860000"
## [1] "At test #: 870000"
## [1] "At test #: 880000"
## [1] "At test #: 890000"
## [1] "At test #: 900000"
## [1] "At test #: 910000"
## [1] "At test #: 920000"
## [1] "At test #: 930000"
## [1] "At test #: 940000"
```

```
## [1] "At test #: 950000"
## [1] "At test #: 960000"
## [1] "At test #: 970000"
## [1] "At test #: 980000"
## [1] "At test #: 990000"
## [1] "At test #: 1000000"
## [1] "At test #: 1010000"
## [1] "At test #: 1020000"
## [1] "At test #: 1030000"
## [1] "At test #: 1040000"
## [1] "At test #: 1050000"
## [1] "At test #: 1060000"
## [1] "At test #: 1070000"
## [1] "At test #: 1080000"
## [1] "At test #: 1090000"
## [1] "At test #: 1100000"
## [1] "At test #: 1110000"
## [1] "At test #: 1120000"
## [1] "At test #: 1130000"
## [1] "At test #: 1140000"
## [1] "At test #: 1150000"
## [1] "At test #: 1160000"
## [1] "At test #: 1170000"
## [1] "At test #: 1180000"
## [1] "At test #: 1190000"
## [1] "At test #: 1200000"
## [1] "At test #: 1210000"
## [1] "At test #: 1220000"
## [1] "At test #: 1230000"
## [1] "At test #: 1240000"
## [1] "At test #: 1250000"
## [1] "At test #: 1260000"
## [1] "At test #: 1270000"
## [1] "At test #: 1280000"
## [1] "At test #: 1290000"
## [1] "At test #: 1300000"
## [1] "At test #: 1310000"
## [1] "At test #: 1320000"
## [1] "At test #: 1330000"
## [1] "At test #: 1340000"
## [1] "At test #: 1350000"
## [1] "At test #: 1360000"
## [1] "At test #: 1370000"
## [1] "At test #: 1380000"
## [1] "At test #: 1390000"
## [1] "At test #: 1400000"
## [1] "At test #: 1410000"
## [1] "At test #: 1420000"
## [1] "At test #: 1430000"
## [1] "At test #: 1440000"
## [1] "At test #: 1450000"
## [1] "At test #: 1460000"
## [1] "At test #: 1470000"
## [1] "At test #: 1480000"
```

```
## [1] "At test #: 1490000"
## [1] "At test #: 1500000"
## [1] "At test #: 1510000"
## [1] "At test #: 1520000"
## [1] "At test #: 1530000"
## [1] "At test #: 1540000"
## [1] "At test #: 1550000"
## [1] "At test #: 1560000"
## [1] "At test #: 1570000"
## [1] "At test #: 1580000"
## [1] "At test #: 1590000"
## [1] "At test #: 1600000"
## [1] "At test #: 1610000"
## [1] "At test #: 1620000"
## [1] "At test #: 1630000"
## [1] "At test #: 1640000"
## [1] "At test #: 1650000"
## [1] "At test #: 1660000"
## [1] "At test #: 1670000"
## [1] "At test #: 1680000"
## [1] "At test #: 1690000"
## [1] "At test #: 1700000"
## [1] "At test #: 1710000"
## [1] "At test #: 1720000"
## [1] "At test #: 1730000"
## [1] "At test #: 1740000"
## [1] "At test #: 1750000"
## [1] "At test #: 1760000"
## [1] "At test #: 1770000"
## [1] "At test #: 1780000"
## [1] "At test #: 1790000"
## [1] "At test #: 1800000"
## [1] "At test #: 1810000"
## [1] "At test #: 1820000"
## [1] "At test #: 1830000"
## [1] "At test #: 1840000"
## [1] "At test #: 1850000"
## [1] "At test #: 1860000"
## [1] "At test #: 1870000"
## [1] "At test #: 1880000"
## [1] "At test #: 1890000"
## [1] "At test #: 1900000"
## [1] "At test #: 1910000"
## [1] "At test #: 1920000"
## [1] "At test #: 1930000"
## [1] "At test #: 1940000"
## [1] "At test #: 1950000"
## [1] "At test #: 1960000"
## [1] "At test #: 1970000"
## [1] "At test #: 1980000"
## [1] "At test #: 1990000"
## [1] "At test #: 2000000"
## [1] "At test #: 2010000"
## [1] "At test #: 2020000"
```

```
## [1] "At test #: 2030000"
## [1] "At test #: 2040000"
## [1] "At test #: 2050000"
## [1] "At test #: 2060000"
## [1] "At test #: 2070000"
## [1] "At test #: 2080000"
## [1] "At test #: 2090000"
## [1] "At test #: 2100000"
## [1] "At test #: 2110000"
## [1] "At test #: 2120000"
## [1] "At test #: 2130000"
## [1] "At test #: 2140000"
## [1] "At test #: 2150000"
## [1] "At test #: 2160000"
## [1] "At test #: 2170000"
## [1] "At test #: 2180000"
## [1] "At test #: 2190000"
## [1] "At test #: 2200000"
## [1] "At test #: 2210000"
## [1] "At test #: 2220000"
## [1] "At test #: 2230000"
## [1] "At test #: 2240000"
## [1] "At test #: 2250000"
## [1] "At test #: 2260000"
## [1] "At test #: 2270000"
## [1] "At test #: 2280000"
## [1] "At test #: 2290000"
## [1] "At test #: 2300000"
## [1] "At test #: 2310000"
## [1] "At test #: 2320000"
## [1] "At test #: 2330000"
## [1] "At test #: 2340000"
## [1] "At test #: 2350000"
## [1] "At test #: 2360000"
## [1] "At test #: 2370000"
## [1] "At test #: 2380000"
## [1] "At test #: 2390000"
## [1] "At test #: 2400000"
## [1] "At test #: 2410000"
## [1] "At test #: 2420000"
## [1] "At test #: 2430000"
## [1] "At test #: 2440000"
## [1] "At test #: 2450000"
## [1] "At test #: 2460000"
## [1] "At test #: 2470000"
## [1] "At test #: 2480000"
## [1] "At test #: 2490000"
## [1] "At test #: 2500000"
```

```
##caution: this should take about an hour!
```

Save results:

```
fitted0.8 <- pi0Est$pi0.lambda[,pi0Est$lambda==0.8]
fitted0.9 <- pi0Est$pi0.lambda[,round(pi0Est$lambda,2)==0.9]
fitted.final.smooth <- pi0Est$pi0

save(fitted0.8, fitted0.9, fitted.final.smooth, file="BMI_GIANT_GWAS_results.RData")
```

Session Information

```
devtools::session_info()
```

```
## Session info -------------------------------------------------------------

##  setting  value
##  version  R version 3.3.1 (2016-06-21)
##  system   x86_64, mingw32
##  ui       RTerm
##  language (EN)
##  collate  English_United States.1252
##  tz       America/New_York
##  date     2016-10-27

## Packages -----------------------------------------------------------------

##  package      * version    date       source
##  acepack        1.3-3.3    2014-11-24 CRAN (R 3.3.0)
##  assertthat     0.1        2013-12-06 CRAN (R 3.3.1)
##  betareg      * 3.0-5      2014-09-25 CRAN (R 3.3.1)
##  BiocStyle    * 2.0.2      2016-05-16 Bioconductor
##  chron          2.3-47     2015-06-24 CRAN (R 3.3.1)
##  cluster        2.0.4      2016-04-18 CRAN (R 3.3.1)
##  codetools      0.2-14     2015-07-15 CRAN (R 3.3.1)
##  colorspace     1.2-6      2015-03-11 CRAN (R 3.3.1)
##  data.table     1.9.6      2015-09-19 CRAN (R 3.3.1)
##  DBI            0.4-1      2016-05-08 CRAN (R 3.3.1)
##  devtools       1.12.0     2016-06-24 CRAN (R 3.3.1)
##  digest         0.6.9      2016-01-08 CRAN (R 3.3.1)
##  dplyr        * 0.5.0      2016-06-24 CRAN (R 3.3.1)
##  evaluate       0.9        2016-04-29 CRAN (R 3.3.1)
##  fdrtool      * 1.2.15     2015-07-08 CRAN (R 3.3.0)
##  flexmix        2.3-13     2015-01-17 CRAN (R 3.3.1)
##  foreign        0.8-66     2015-08-19 CRAN (R 3.3.1)
##  formatR        1.4        2016-05-09 CRAN (R 3.3.1)
##  Formula      * 1.2-1      2015-04-07 CRAN (R 3.3.0)
##  ggplot2      * 2.1.0      2016-03-01 CRAN (R 3.3.1)
##  gridExtra      2.2.1      2016-02-29 CRAN (R 3.3.1)
##  gtable         0.2.0      2016-02-26 CRAN (R 3.3.1)
##  Hmisc        * 3.17-4     2016-05-02 CRAN (R 3.3.1)
##  htmltools      0.3.5      2016-03-21 CRAN (R 3.3.1)
##  knitr          1.13       2016-05-09 CRAN (R 3.3.1)
##  lattice      * 0.20-33    2015-07-14 CRAN (R 3.3.1)
```

```
##  latticeExtra   0.6-28     2016-02-09 CRAN (R 3.3.1)
##  lmtest         0.9-34     2015-06-06 CRAN (R 3.3.1)
##  magrittr       1.5        2014-11-22 CRAN (R 3.3.1)
##  Matrix         1.2-6      2016-05-02 CRAN (R 3.3.1)
##  memoise        1.0.0      2016-01-29 CRAN (R 3.3.1)
##  modeltools     0.2-21     2013-09-02 CRAN (R 3.3.0)
##  munsell        0.4.3      2016-02-13 CRAN (R 3.3.1)
##  nnet           7.3-12     2016-02-02 CRAN (R 3.3.1)
##  plyr           1.8.4      2016-06-08 CRAN (R 3.3.1)
##  R6             2.1.2      2016-01-26 CRAN (R 3.3.1)
##  RColorBrewer   1.1-2      2014-12-07 CRAN (R 3.3.0)
##  Rcpp           0.12.6     2016-07-19 CRAN (R 3.3.1)
##  readr        * 0.2.2      2015-10-22 CRAN (R 3.3.1)
##  reshape2     * 1.4.1      2014-12-06 CRAN (R 3.3.1)
##  rmarkdown      1.0        2016-07-08 CRAN (R 3.3.1)
##  rpart          4.1-10     2015-06-29 CRAN (R 3.3.1)
##  sandwich       2.3-4      2015-09-24 CRAN (R 3.3.1)
##  scales         0.4.0      2016-02-26 CRAN (R 3.3.1)
##  stringi        1.1.1      2016-05-27 CRAN (R 3.3.0)
##  stringr        1.0.0      2015-04-30 CRAN (R 3.3.1)
##  survival     * 2.39-4     2016-05-11 CRAN (R 3.3.1)
##  swfdr        * 0.0.0.9000 2016-10-21 local
##  tibble         1.1        2016-07-04 CRAN (R 3.3.1)
##  withr          1.0.2      2016-06-20 CRAN (R 3.3.1)
##  yaml           2.1.13     2014-06-12 CRAN (R 3.3.1)
##  zoo            1.7-13     2016-05-03 CRAN (R 3.3.1)
```