

Analysis of BMI GIANT GWAS data

Simina Boca, Jeff Leek

2017-06-05

Load the relevant libraries:

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

source("../functions.R")
```

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 without any covariates

```
qVals <- qvalue(tot$p)
cbind(qVals$lambda, qVals$pi0.lambda)
```

```
##      [,1]      [,2]
## [1,] 0.05 0.9705014
## [2,] 0.10 0.9651139
## [3,] 0.15 0.9624994
## [4,] 0.20 0.9608043
## [5,] 0.25 0.9590810
## [6,] 0.30 0.9576285
## [7,] 0.35 0.9561064
## [8,] 0.40 0.9556923
## [9,] 0.45 0.9537792
## [10,] 0.50 0.9543245
## [11,] 0.55 0.9547936
## [12,] 0.60 0.9508101
## [13,] 0.65 0.9513008
## [14,] 0.70 0.9533961
## [15,] 0.75 0.9496255
## [16,] 0.80 0.9513779
## [17,] 0.85 0.9397846
## [18,] 0.90 0.9493664
## [19,] 0.95 0.9551331
```

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)
```

```
## At test #: 10000
## At test #: 20000
## At test #: 30000
## At test #: 40000
## At test #: 50000
## At test #: 60000
## At test #: 70000
## At test #: 80000
## At test #: 90000
## At test #: 100000
## At test #: 110000
## At test #: 120000
```

At test #: 130000
At test #: 140000
At test #: 150000
At test #: 160000
At test #: 170000
At test #: 180000
At test #: 190000
At test #: 200000
At test #: 210000
At test #: 220000
At test #: 230000
At test #: 240000
At test #: 250000
At test #: 260000
At test #: 270000
At test #: 280000
At test #: 290000
At test #: 300000
At test #: 310000
At test #: 320000
At test #: 330000
At test #: 340000
At test #: 350000
At test #: 360000
At test #: 370000
At test #: 380000
At test #: 390000
At test #: 400000
At test #: 410000
At test #: 420000
At test #: 430000
At test #: 440000
At test #: 450000
At test #: 460000
At test #: 470000
At test #: 480000

At test #: 490000
At test #: 500000
At test #: 510000
At test #: 520000
At test #: 530000
At test #: 540000
At test #: 550000
At test #: 560000
At test #: 570000
At test #: 580000
At test #: 590000
At test #: 600000
At test #: 610000
At test #: 620000
At test #: 630000
At test #: 640000
At test #: 650000
At test #: 660000
At test #: 670000
At test #: 680000
At test #: 690000
At test #: 700000
At test #: 710000
At test #: 720000
At test #: 730000
At test #: 740000
At test #: 750000
At test #: 760000
At test #: 770000
At test #: 780000
At test #: 790000
At test #: 800000
At test #: 810000
At test #: 820000
At test #: 830000
At test #: 840000

At test #: 850000
At test #: 860000
At test #: 870000
At test #: 880000
At test #: 890000
At test #: 900000
At test #: 910000
At test #: 920000
At test #: 930000
At test #: 940000
At test #: 950000
At test #: 960000
At test #: 970000
At test #: 980000
At test #: 990000
At test #: 1000000
At test #: 1010000
At test #: 1020000
At test #: 1030000
At test #: 1040000
At test #: 1050000
At test #: 1060000
At test #: 1070000
At test #: 1080000
At test #: 1090000
At test #: 1100000
At test #: 1110000
At test #: 1120000
At test #: 1130000
At test #: 1140000
At test #: 1150000
At test #: 1160000
At test #: 1170000
At test #: 1180000
At test #: 1190000
At test #: 1200000

At test #: 1210000
At test #: 1220000
At test #: 1230000
At test #: 1240000
At test #: 1250000
At test #: 1260000
At test #: 1270000
At test #: 1280000
At test #: 1290000
At test #: 1300000
At test #: 1310000
At test #: 1320000
At test #: 1330000
At test #: 1340000
At test #: 1350000
At test #: 1360000
At test #: 1370000
At test #: 1380000
At test #: 1390000
At test #: 1400000
At test #: 1410000
At test #: 1420000
At test #: 1430000
At test #: 1440000
At test #: 1450000
At test #: 1460000
At test #: 1470000
At test #: 1480000
At test #: 1490000
At test #: 1500000
At test #: 1510000
At test #: 1520000
At test #: 1530000
At test #: 1540000
At test #: 1550000
At test #: 1560000

At test #: 1570000
At test #: 1580000
At test #: 1590000
At test #: 1600000
At test #: 1610000
At test #: 1620000
At test #: 1630000
At test #: 1640000
At test #: 1650000
At test #: 1660000
At test #: 1670000
At test #: 1680000
At test #: 1690000
At test #: 1700000
At test #: 1710000
At test #: 1720000
At test #: 1730000
At test #: 1740000
At test #: 1750000
At test #: 1760000
At test #: 1770000
At test #: 1780000
At test #: 1790000
At test #: 1800000
At test #: 1810000
At test #: 1820000
At test #: 1830000
At test #: 1840000
At test #: 1850000
At test #: 1860000
At test #: 1870000
At test #: 1880000
At test #: 1890000
At test #: 1900000
At test #: 1910000
At test #: 1920000

At test #: 1930000
At test #: 1940000
At test #: 1950000
At test #: 1960000
At test #: 1970000
At test #: 1980000
At test #: 1990000
At test #: 2000000
At test #: 2010000
At test #: 2020000
At test #: 2030000
At test #: 2040000
At test #: 2050000
At test #: 2060000
At test #: 2070000
At test #: 2080000
At test #: 2090000
At test #: 2100000
At test #: 2110000
At test #: 2120000
At test #: 2130000
At test #: 2140000
At test #: 2150000
At test #: 2160000
At test #: 2170000
At test #: 2180000
At test #: 2190000
At test #: 2200000
At test #: 2210000
At test #: 2220000
At test #: 2230000
At test #: 2240000
At test #: 2250000
At test #: 2260000
At test #: 2270000
At test #: 2280000


```
## At test #: 2290000
## At test #: 2300000
## At test #: 2310000
## At test #: 2320000
## At test #: 2330000
## At test #: 2340000
## At test #: 2350000
## At test #: 2360000
## At test #: 2370000
## At test #: 2380000
## At test #: 2390000
## At test #: 2400000
## At test #: 2410000
## At test #: 2420000
## At test #: 2430000
## At test #: 2440000
## At test #: 2450000
## At test #: 2460000
## At test #: 2470000
## At test #: 2480000
## At test #: 2490000
## 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_logistic.RData")
```

Session Information

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

```
## Session info -----
## setting value
## version R version 3.4.0 (2017-04-21)
## system x86_64, mingw32
## ui RTerm
## language (EN)
```

```
## collate English_United States.1252
## tz      America/New_York
## date    2017-06-05
```

```
## Packages -----
```

## package	* version	date	source
## acepack	1.4.1	2016-10-29	CRAN (R 3.4.0)
## assertthat	0.2.0	2017-04-11	CRAN (R 3.4.0)
## backports	1.0.5	2017-01-18	CRAN (R 3.4.0)
## base64enc	0.1-3	2015-07-28	CRAN (R 3.4.0)
## betareg	* 3.1-0	2016-08-06	CRAN (R 3.4.0)
## BiocStyle	* 2.4.0	2017-04-25	Bioconductor
## checkmate	1.8.2	2016-11-02	CRAN (R 3.4.0)
## cluster	2.0.6	2017-03-10	CRAN (R 3.4.0)
## colorspace	1.3-2	2016-12-14	CRAN (R 3.4.0)
## data.table	1.10.4	2017-02-01	CRAN (R 3.4.0)
## DBI	0.6-1	2017-04-01	CRAN (R 3.4.0)
## devtools	1.12.0	2016-12-05	CRAN (R 3.4.0)
## digest	0.6.12	2017-01-27	CRAN (R 3.4.0)
## dplyr	* 0.5.0	2016-06-24	CRAN (R 3.4.0)
## evaluate	0.10	2016-10-11	CRAN (R 3.4.0)
## fdrtol	* 1.2.15	2015-07-08	CRAN (R 3.4.0)
## flexmix	2.3-14	2017-04-28	CRAN (R 3.4.0)
## foreign	0.8-67	2016-09-13	CRAN (R 3.4.0)
## Formula	* 1.2-1	2015-04-07	CRAN (R 3.4.0)
## ggplot2	* 2.2.1	2016-12-30	CRAN (R 3.4.0)
## gridExtra	2.2.1	2016-02-29	CRAN (R 3.4.0)
## gtable	0.2.0	2016-02-26	CRAN (R 3.4.0)
## Hmisc	* 4.0-3	2017-05-02	CRAN (R 3.4.0)
## hms	0.3	2016-11-22	CRAN (R 3.4.0)
## htmlTable	1.9	2017-01-26	CRAN (R 3.4.0)
## htmltools	0.3.5	2016-03-21	CRAN (R 3.4.0)
## htmlwidgets	0.8	2016-11-09	CRAN (R 3.4.0)
## knitr	1.15.1	2016-11-22	CRAN (R 3.4.0)
## lattice	* 0.20-35	2017-03-25	CRAN (R 3.4.0)
## latticeExtra	0.6-28	2016-02-09	CRAN (R 3.4.0)
## lazyeval	0.2.0	2016-06-12	CRAN (R 3.4.0)
## lmtest	0.9-35	2017-02-11	CRAN (R 3.4.0)
## magrittr	1.5	2014-11-22	CRAN (R 3.4.0)
## Matrix	1.2-9	2017-03-14	CRAN (R 3.4.0)
## memoise	1.1.0	2017-04-21	CRAN (R 3.4.0)
## modeltools	0.2-21	2013-09-02	CRAN (R 3.4.0)
## munsell	0.4.3	2016-02-13	CRAN (R 3.4.0)
## nnet	7.3-12	2016-02-02	CRAN (R 3.4.0)
## plyr	1.8.4	2016-06-08	CRAN (R 3.4.0)
## qvalue	* 2.8.0	2017-04-25	Bioconductor
## R6	2.2.0	2016-10-05	CRAN (R 3.4.0)
## RColorBrewer	1.1-2	2014-12-07	CRAN (R 3.4.0)
## Rcpp	0.12.10	2017-03-19	CRAN (R 3.4.0)
## readr	* 1.1.0	2017-03-22	CRAN (R 3.4.0)
## reshape2	* 1.4.2	2016-10-22	CRAN (R 3.4.0)
## rmarkdown	1.5	2017-04-26	CRAN (R 3.4.0)
## rpart	4.1-11	2017-03-13	CRAN (R 3.4.0)
## rprojroot	1.2	2017-01-16	CRAN (R 3.4.0)

```
## sandwich      2.3-4    2015-09-24 CRAN (R 3.4.0)
## scales        0.4.1    2016-11-09 CRAN (R 3.4.0)
## stringi       1.1.5    2017-04-07 CRAN (R 3.4.0)
## stringr       1.2.0    2017-02-18 CRAN (R 3.4.0)
## survival      * 2.41-3  2017-04-04 CRAN (R 3.4.0)
## tibble        1.3.0    2017-04-01 CRAN (R 3.4.0)
## withr         1.0.2    2016-06-20 CRAN (R 3.4.0)
## yaml          2.1.14   2016-11-12 CRAN (R 3.4.0)
## zoo           1.8-0    2017-04-12 CRAN (R 3.4.0)
```