/home/nicole/Jupyter/JG3/Data/0.5/M/2

```
;ls
In [4]:
        Correlation.G5.M.JC.txt
        G0.Genotype.ID
        G0.ID
        G0.noGenotype.ID
        G1.Genotype.ID
        G1.ID
        G1.noGenotype.ID
        G2.Genotype.ID
        G2.ID
        G2.noGenotype.ID
        G3.Genotype.ID
        G3.ID
        G3.noGenotype.ID
        G4.Genotype.ID
        G4.ID
        G4.noGenotype.ID
        G5.Genotype.ID
        G5.ID
        G5.noGenotype.ID
        MarNF.txt
        PedAll.txt
        Phe.txt
        PheAll.txt
        Regression.G5.M.JC.txt
        all.ID
        alphaEstimatesJ
        alphaEstimatesJC
        epsiEstimatesJ
        epsiEstimatesJC
        genotype.ID
        meanOfSNPMAll
        meanOfSNPMG0
        meanOfSNPMG1
        meanOfSNPMG2
        meanOfSNPMG3
        meanOfSNPMG4
        meanOfSNPMG5
        noGenotype.ID
        sim.bv
        sim.phenotype
In [5]: ;awk '{print $1}' PedAll.txt | sort -b > all.ID
In [6]:
        ;awk '{print $1}' MarNF.txt | sort -b > genotype.ID
In [7]:
        ;join -v1 all.ID genotype.ID > noGenotype.ID
In [8]:
        ;awk '{print $1,$2}' Phe.txt > sim.phenotype
In [9]: ;awk '{print $1,$3}' PheAll.txt > sim.bv
```

```
In [10]: ; awk 'NR >=1 && NR <=8000 {print $1}' PedAll.txt | sort -b > G0.ID
In [11]: ; awk 'NR >=8001 && NR <=16000 {print $1}' PedAll.txt | sort -b > G1.ID
         ; awk 'NR >=16001 && NR <=24000 {print $1}' PedAll.txt | sort -b > G2.ID
In [12]:
         ; awk 'NR >=24001 && NR <=32000 {print $1}' PedAll.txt | sort -b > G3.ID
         ; awk 'NR >=32001 && NR <=40000 {print $1}' PedAll.txt | sort -b > G4.ID
In [15]:
         ; awk 'NR >=40001 && NR <=48000 {print $1}' PedAll.txt | sort -b > G5.ID
         ; join G0.ID genotype.ID > G0.Genotype.ID
In [16]:
In [17]:
         ; join G1.ID genotype.ID > G1.Genotype.ID
In [18]:
         ; join G2.ID genotype.ID > G2.Genotype.ID
In [19]:
         ;join G3.ID genotype.ID > G3.Genotype.ID
In [20]:
         ; join G4.ID genotype.ID > G4.Genotype.ID
         ; join G5.ID genotype.ID > G5.Genotype.ID
In [21]:
In [22]:
         ; join -v1 G0.ID genotype.ID > G0.noGenotype.ID
In [23]:
         ; join -v1 G1.ID genotype.ID > G1.noGenotype.ID
In [24]:
         ; join -v1 G2.ID genotype.ID > G2.noGenotype.ID
         ;join -v1 G3.ID genotype.ID > G3.noGenotype.ID
In [26]:
         ; join -v1 G4.ID genotype.ID > G4.noGenotype.ID
In [27]:
         ; join -v1 G5.ID genotype.ID > G5.noGenotype.ID
In [28]:
         ;wc G0.Genotype.ID;wc G1.Genotype.ID;wc G2.Genotype.ID;wc G3.Genotype.ID;wc (
          200
               200 1200 GO.Genotype.ID
          200
               200 1200 G1.Genotype.ID
               200 1200 G2.Genotype.ID
          200
              200 1200 G3.Genotype.ID
               200 1200 G4.Genotype.ID
          200
          8000 8000 48000 G5.Genotype.ID
```

```
;wc G0.noGenotype.ID;wc G1.noGenotype.ID;wc G2.noGenotype.ID;wc G3.noGenotype
In [29]:
                7800 46800 G0.noGenotype.ID
          7800
                7800 46800 Gl.noGenotype.ID
          7800 7800 46800 G2.noGenotype.ID
          7800 7800 46800 G3.noGenotype.ID
          7800 7800 46800 G4.noGenotype.ID
         0 0 0 G5.noGenotype.ID
         ped,A Mats,numSSBayes = calc Ai("PedAll.txt", "genotype.ID", calculateInbreedia
In [30]:
         nothing
         df
                 = read_genotypes("MarNF.txt",numSSBayes)
         M Mats = make MMats(df, A Mats, ped, center=true);
                                                                                  # with
         y_Vecs = make_yVecs("sim.phenotype",ped,numSSBayes);
         Z_Mats = make_ZMats(ped,y_Vecs,numSSBayes)
         X Mats, W Mats = make XWMats(Z Mats, M Mats, numSSBayes)
                                                                                  # no
         nothing
In [31]: | vRes
                = 0.714
                = 0.714
         vG
         nIter = 50000
         @time aHat1,alphaHat,betaHat,epsiHat =
         ssGibbs(M_Mats,y_Vecs,Z_Mats,X_Mats,W_Mats,A_Mats, numSSBayes,vRes,vG,nIter,
         nothing
         This is iteration 5000
         This is iteration 10000
         This is iteration 15000
         This is iteration 20000
         This is iteration 25000
         This is iteration 30000
         This is iteration 35000
         This is iteration 40000
         This is iteration 45000
         This is iteration 50000
         2388.704963 seconds (23.04 G allocations: 723.736 GB, 7.46% gc time)
In [32]: | betaHat
Out[32]: 1-element Array{Float64,1}:
          10.5168
```

```
alphaHat
In [33]:
Out[33]: 150-element Array{Float64,1}:
          -0.0511688
          -0.115607
           0.196472
           0.0247497
          -0.0285359
            0.0988894
          -0.0885265
           0.0506333
          -0.0691985
           0.034203
          -0.0599119
          -0.0296326
           0.0382986
          -0.12435
           0.129464
          -0.1103
          -0.0529531
           0.165205
           0.0296973
          -0.148262
           0.034057
           0.0424689
            0.0699243
           -0.0499534
            0.0605296
In [34]: writedlm("alphaEstimatesC",alphaHat)
```

```
In [35]: epsiHat
Out[35]: 45930-element Array{Float64,1}:
           0.483664
           0.193591
           0.299089
           0.251221
          -0.164474
           0.296757
          -0.214947
           0.196153
          -0.205759
          -0.0445675
           0.409274
          -0.166966
          -0.366344
          -0.450971
          -0.000529153
          -0.788038
           0.100398
          -0.766421
          -0.600181
          -0.152897
           0.491772
           0.517489
          -0.12128
          -0.736403
           0.654348
In [36]: writedlm("epsiEstimatesC",epsiHat)
In [37]: using DataFrames
In [38]: df = readtable("sim.bv", eltypes =[UTF8String, Float64], separator = ' ',head
         a = Array(Float64, numSSBayes.num ped)
         for (i,ID) in enumerate(df[:,1])
             j = ped.idMap[ID].seqID
             a[j] = df[i,2]
         end
In [39]: IDs = readtable("all.ID", eltypes =[UTF8String], separator = ' ',header=false
         posAi = getPos(ped,IDs)
         cor1 = cor(a[posAi],aHat1[posAi])[1,1]
         reg1 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - all.ID : correlation = %6.3f\n", cor1 ) # with 
         @printf("SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = %6.3f\n",
         JCAll = cor1
         SSBRJC from Gibbs - all.ID : correlation = 0.898
         SSBRJC from Gibbs - all.ID : regression of TBV on GEBV = 0.997
Out[39]: 0.8980693194548445
```

```
In [40]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[40]: 0.1398402531135769
In [41]: IDs = readtable("genotype.ID", eltypes =[UTF8String], separator = ' ',header:
         posAi = getPos(ped,IDs)
         cor2 = cor(a[posAi],aHat1[posAi])[1,1]
         reg2 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - genotype.ID : correlation = %6.3f\n", cor2 ) # |
         @printf("SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = %6.3:
         JCAll = cor2
         SSBRJC from Gibbs - genotype.ID : correlation = 0.802
         SSBRJC from Gibbs - genotype.ID : regression of TBV on GEBV = 0.967
Out[41]: 0.8024121033191799
In [42]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[42]: 1.2477165880879517
In [43]: IDs = readtable("noGenotype.ID", eltypes =[UTF8String], separator = ' ',heade
         posAi = getPos(ped,IDs)
         cor3 = cor(a[posAi],aHat1[posAi])[1,1]
         reg3 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - noGenotype.ID : correlation = %6.3f\n", cor3 );
         @printf("SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = %6
         JCAll = cor3
         SSBRJC from Gibbs - noGenotype.ID : correlation = 0.870
         SSBRJC from Gibbs - noGenotype.ID : regression of TBV on GEBV = 0.966
Out[43]: 0.8704994485251227
In [44]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[44]: -0.11582351649589426
In [45]: IDs = readtable("G0.ID", eltypes =[UTF8String], separator = ' ',header=false
         posAi = getPos(ped,IDs)
         cor4 = cor(a[posAi],aHat1[posAi])[1,1]
         reg4 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G0.ID : correlation = %6.3f\n", cor4 ) # with ep
         @printf("SSBRJC from Gibbs - G0.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor4
         SSBRJC from Gibbs - G0.ID : correlation = 0.709
         SSBRJC from Gibbs - G0.ID: regression of TBV on GEBV = 0.860
Out[45]: 0.7093541737528876
```

```
In [46]: | GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[46]: -1.267547271562283
In [47]: IDs = readtable("G1.ID", eltypes =[UTF8String], separator = ' ',header=false
         posAi = getPos(ped,IDs)
         cor4 = cor(a[posAi],aHat1[posAi])[1,1]
         reg4 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G1.ID : correlation = %6.3f\n", cor4 ) # with e;
         @printf("SSBRJC from Gibbs - G1.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor4
         SSBRJC from Gibbs - G1.ID : correlation =
         SSBRJC from Gibbs - G1.ID: regression of TBV on GEBV = 0.983
Out[47]: 0.76051915558157
In [48]: GEBV = aHat1[posAi]
         G1GEBV=mean(GEBV)
Out[48]: -0.5603515446151832
In [49]: IDs = readtable("G2.ID", eltypes =[UTF8String], separator = ' ', header=false
         posAi = getPos(ped,IDs)
         cor5 = cor(a[posAi],aHat1[posAi])[1,1]
         req5 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G2.ID : correlation = %6.3f\n", cor5 ) # with ei
         @printf("SSBRJC from Gibbs - G2.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor5
         SSBRJC from Gibbs - G2.ID : correlation = 0.749
         SSBRJC from Gibbs - G2.ID: regression of TBV on GEBV = 0.996
Out[49]: 0.7491694765939771
In [50]: GEBV = aHat1[posAi]
         G2GEBV=mean(GEBV)
Out[50]: -0.006607315209380545
In [51]: IDs = readtable("G3.ID", eltypes =[UTF8String], separator = ' ',header=false
         posAi = getPos(ped,IDs)
         cor6 = cor(a[posAi],aHat1[posAi])[1,1]
         reg6 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G3.ID : correlation = %6.3f\n", cor6 ) # with ei
         @printf("SSBRJC from Gibbs - G3.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor6
         SSBRJC from Gibbs - G3.ID : correlation = 0.733
         SSBRJC from Gibbs - G3.ID: regression of TBV on GEBV = 0.973
Out[51]: 0.733196324600902
```

```
In [52]: | GEBV = aHat1[posAi]
         G3GEBV=mean(GEBV)
Out[52]: 0.46322660700915935
In [53]: IDs = readtable("G4.ID", eltypes =[UTF8String], separator = ' ',header=false
         posAi = getPos(ped,IDs)
         cor7 = cor(a[posAi],aHat1[posAi])[1,1]
         reg7 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G4.ID : correlation = %6.3f\n", cor7 ) # with e;
         @printf("SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor7
         SSBRJC from Gibbs - G4.ID : correlation =
         SSBRJC from Gibbs - G4.ID : regression of TBV on GEBV = 0.998
Out[53]: 0.7409451738310643
In [54]: GEBV = aHat1[posAi]
         G4GEBV=mean(GEBV)
Out[54]: 0.9267791928911561
In [55]: IDs = readtable("G5.ID", eltypes =[UTF8String], separator = ' ',header=false
         posAi = getPos(ped,IDs)
         cor8 = cor(a[posAi],aHat1[posAi])[1,1]
         reg8 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G5.ID : correlation = %6.3f\n", cor8 ) # with ep
         @printf("SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = %6.3f\n", re
         JCAll = cor8
         SSBRJC from Gibbs - G5.ID : correlation = 0.778
         SSBRJC from Gibbs - G5.ID : regression of TBV on GEBV = 0.923
Out[55]: 0.7782186546304508
In [56]: GEBV = aHat1[posAi]
         G5GEBV=mean(GEBV)
Out[56]: 1.283541850167993
In [57]: GEBVG5G1=G5GEBV-G1GEBV
Out[57]: 1.8438933947831762
In [58]: GEBVG1G5=[G1GEBV;G2GEBV;G3GEBV;G4GEBV;G5GEBV]
Out[58]: 5-element Array{Float64,1}:
          -0.560352
          -0.00660732
           0.463227
           0.926779
           1.28354
```

```
In [59]: reg8 = linreg(aHat1[posAi], a[posAi])
Out[59]: 2-element Array{Float64,1}:
          10.7525
           0.923
In [60]: VarGEBV=var(aHat1[posAi])
Out[60]: 0.33889528347017783
In [61]: VarTBV=var(a[posAi])
Out[61]: 0.4767229049180523
In [62]: Cov=cov(aHat1[posAi], a[posAi])
Out[62]: 0.31280047851535403
In [63]: b=Cov/VarGEBV
Out[63]: 0.9230003891242703
In [64]: IDs = readtable("G0.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         posAi = getPos(ped,IDs)
         cor9 = cor(a[posAi],aHat1[posAi])[1,1]
         reg9 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G0.Genotype.ID : correlation = %6.3f\n", cor9 )
         @printf("SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = %6
         JCAll = cor9
         SSBRJC from Gibbs - G0.Genotype.ID : correlation = 0.805
         SSBRJC from Gibbs - G0.Genotype.ID : regression of TBV on GEBV = 0.999
Out[64]: 0.8048951116214481
In [65]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[65]: 0.22312107123050529
In [66]: IDs = readtable("G1.Genotype.ID", eltypes =[UTF8String], separator = ' ', head
         posAi = getPos(ped,IDs)
         cor9 = cor(a[posAi],aHat1[posAi])[1,1]
         reg9 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G1.Genotype.ID : correlation = %6.3f\n", cor9 )
         @printf("SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = %6
         JCAll = cor9
         SSBRJC from Gibbs - G1.Genotype.ID : correlation =
                                                              0.765
         SSBRJC from Gibbs - G1.Genotype.ID : regression of TBV on GEBV = 0.858
Out[66]: 0.7646228117343004
```

```
In [67]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[67]: 0.6370612189809444
In [68]: IDs = readtable("G2.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         posAi = getPos(ped,IDs)
         cor10 = cor(a[posAi],aHat1[posAi])[1,1]
         reg10 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G2.Genotype.ID : correlation = %6.3f\n", cor10
         @printf("SSBRJC from Gibbs - G2.Genotype.ID : regression of TBV on GEBV = %6
         JCAll = cor10
         SSBRJC from Gibbs - G2.Genotype.ID : correlation = 0.766
         SSBRJC from Gibbs - G2.Genotype.ID: regression of TBV on GEBV = 0.864
Out[68]: 0.7655191056483671
In [69]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[69]: 0.9732835931759004
In [70]: IDs = readtable("G3.Genotype.ID", eltypes =[UTF8String], separator = ' ',head
         posAi = getPos(ped,IDs)
         cor11 = cor(a[posAi],aHat1[posAi])[1,1]
         reg11 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G3.Genotype.ID : correlation = %6.3f\n", cor11
         @printf("SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = %6
         JCAll = cor11
         SSBRJC from Gibbs - G3.Genotype.ID : correlation = 0.765
         SSBRJC from Gibbs - G3.Genotype.ID : regression of TBV on GEBV = 0.795
Out[70]: 0.764854333045188
In [71]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[71]: 1.3843006559470064
In [72]: IDs = readtable("G4.Genotype.ID", eltypes =[UTF8String], separator = ' ', head
         posAi = getPos(ped,IDs)
         cor12 = cor(a[posAi],aHat1[posAi])[1,1]
         reg12 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G4.Genotype.ID : correlation = %6.3f\n", cor12
         @printf("SSBRJC from Gibbs - G4.Genotype.ID : regression of TBV on GEBV = %6
         JCAll = cor12
         SSBRJC from Gibbs - G4.Genotype.ID : correlation = 0.752
         SSBRJC from Gibbs - G4.Genotype.ID: regression of TBV on GEBV = 0.782
Out[72]: 0.7524040460517722
```

```
In [73]: | GEBV = aHat1[posAi]
         mean(GEBV)
Out[73]: 1.5878059179037436
In [74]: IDs = readtable("G5.Genotype.ID", eltypes =[UTF8String], separator = ' ',heac
         posAi = getPos(ped,IDs)
         cor13 = cor(a[posAi],aHat1[posAi])[1,1]
         reg13 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G5.Genotype.ID : correlation = %6.3f\n", cor13
         @printf("SSBRJC from Gibbs - G5.Genotype.ID : regression of TBV on GEBV = %6
         JCAll = cor13
         SSBRJC from Gibbs - G5.Genotype.ID : correlation =
         SSBRJC from Gibbs - G5.Genotype.ID: regression of TBV on GEBV = 0.923
Out[74]: 0.7782186546304508
In [75]: writedlm("Correlation.G5.M.C.txt",cor13)
In [76]: | writedlm("Regression.G5.M.C.txt",reg13)
In [77]: GEBVG5Gall = aHat1[posAi]
         GEBVG5G=mean(GEBVG5Gall)
Out[77]: 1.283541850167993
In [78]: IDs = readtable("G0.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         posAi = getPos(ped,IDs)
         cor14 = cor(a[posAi],aHat1[posAi])[1,1]
         reg14 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G0.noGenotype.ID : correlation = %6.3f\n", corl
         @printf("SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = '
         JCAll = cor14
         SSBRJC from Gibbs - G0.noGenotype.ID : correlation = 0.693
         SSBRJC from Gibbs - G0.noGenotype.ID : regression of TBV on GEBV = 0.874
Out[78]: 0.6925079155743108
In [79]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[79]: -1.3057695367620983
```

```
In [80]: IDs = readtable("G1.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         posAi = getPos(ped,IDs)
         cor14 = cor(a[posAi],aHat1[posAi])[1,1]
         reg14 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G1.noGenotype.ID : correlation = %6.3f\n", corl
         @printf("SSBRJC from Gibbs - G1.noGenotype.ID : regression of TBV on GEBV = '
         JCAll = cor14
         SSBRJC from Gibbs - G1.noGenotype.ID : correlation = 0.745
         SSBRJC from Gibbs - G1.noGenotype.ID: regression of TBV on GEBV = 0.990
Out[80]: 0.7454271331903105
In [81]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[81]: -0.5910544359894429
In [82]: IDs = readtable("G2.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         posAi = getPos(ped,IDs)
         cor15 = cor(a[posAi],aHat1[posAi])[1,1]
         reg15 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G2.noGenotype.ID : correlation = %6.3f\n", cor1!
         @printf("SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV = '
         JCAll = cor15
         SSBRJC from Gibbs - G2.noGenotype.ID : correlation = 0.734
         SSBRJC from Gibbs - G2.noGenotype.ID : regression of TBV on GEBV = 0.993
Out[82]: 0.7339164268571134
In [83]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[83]: -0.031732723116695415
In [84]: IDs = readtable("G3.noGenotype.ID", eltypes =[UTF8String], separator = ' ',he
         posAi = getPos(ped,IDs)
         cor16 = cor(a[posAi],aHat1[posAi])[1,1]
         reg16 = linreg(aHat1[posAi], a[posAi])[2,1]
         @printf("SSBRJC from Gibbs - G3.noGenotype.ID : correlation = %6.3f\n", cor1
         @printf("SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = '
         JCAll = cor16
         SSBRJC from Gibbs - G3.noGenotype.ID : correlation = 0.716
         SSBRJC from Gibbs - G3.noGenotype.ID : regression of TBV on GEBV = 0.965
Out[84]: 0.7161503650737407
In [85]: GEBV = aHat1[posAi]
         mean(GEBV)
Out[85]: 0.4396093237030608
```