Calculating LD between biallelic and multiallelic markers with confidence intervals

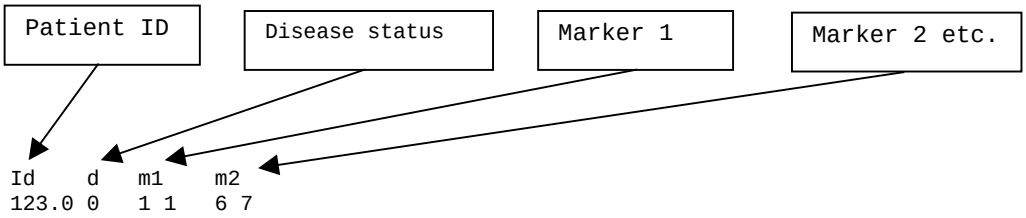
See

http://www.iop.kcl.ac.uk/IoP/Departments/PsychMed/GEpiBSt/software.stm

for the programs

Running PM (or PM+ for using EH+)

The ms-dos input file (example.dat) has the following format:




```
Id    d   m1    m2
123.0 0   1 1   6 7
```

A real example of the file will actually look like this:

```
1003.0 1   1 2   1 2
1003.1 0   1 2   1 2
1003.2 0   1 2   1 1
1005.0 1   1 1   1 2
1005.1 0   1 1   1 2
1005.2 0   1 2   1 1
1006.0 1   2 2   2 2
1006.1 0   1 2   2 2
1006.2 0   2 2   1 2
1007.0 1   1 1   1 2
1007.1 0   1 1   1 2
1007.2 0   1 2   1 1
```

The parameter file (example.par) has the following format

```
2 0 0 0   [no. of loci; case control? 0,1; permutations 0,1; number permutations 0,1]
2 10 [number alleles locus 1; number of allele locus 2]
0 0 [marker phenotype format 0 is two columns, 1 is one column i.e. 1, 2, 3]
1 1 [selection status i.e. which columns to use in the analysis]
0 0 [marker permutations]
0 0 0 0 0 [disease model; disease allele freq.;  3 columns for penetrance]
```

The file will actually look like this:

```
2  1  0  0   << nloci, case/control, label permutation, # permutations
2 2          << a list of marker alleles
0 0          << allele/genotype, screen output
1 1          << marker selection status
0 0          << marker permutation status
0.001 0.05 0.2 0.8 << disease model for case-control design
```

Then run PM using these files to generate input for EH or EH+

>PM example.par example.dat example.out

This creates the file eh.sav (or case.sav and control.sav for a case control study)

Running EH

>EH EH.sav (or EHplus EHplus.sav)

This creates the file EH.out (or EHplus.out):

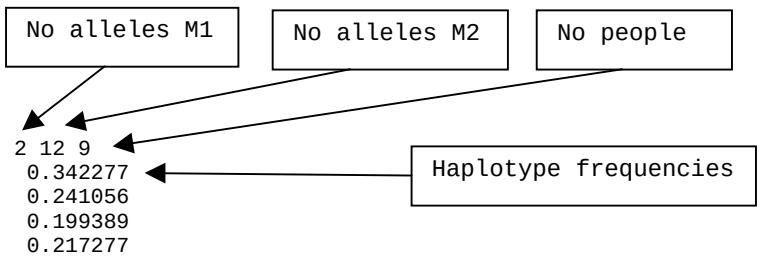
```
Estimates of Gene Frequencies (Assuming Independence)
----\---------------------------
locus \ allele     1       2
--------\------------------------
   1    |         0.5833  0.4167
   2    |         0.5417  0.4583
--------------------------------
# of Typed Individuals: 12

There are 4 Possible Haplotypes of These 2 Loci.
They are Listed Below, with their Estimated Frequencies:
-----------------------------------------------------
| Allele   Allele  |      Haplotype Frequency      |
|   at       at    |                               |
| Locus 1  Locus 2 |  Independent   w/Association  |
-----------------------------------------------------
   1        1           0.315972     0.342277
   1        2           0.267361     0.241056
   2        1           0.225694     0.199389
   2        2           0.190972     0.217277
-----------------------------------------------------
# of Iterations = 7

                                        df   Ln(L)     Chi-square
-------------------------------------------------------------------------
H0: No Association                       2    -23.84      0.00
H1: Allelic Associations Allowed         3    -23.75      0.18
```

Running 2ld

Cut and paste the second **bold** column (haplotype frequencies) into the input file for the program 2LD.



Then run 2ld using this file as the input file.

>2ld example.2ld

Output of 2LD:

```
Estimated haplotype frequencies

         B1       B2
A1 0.042817 0.676695 | 0.719512
A2 0.208707 0.071780 | 0.280487
   ------------------+---------
   0.251524 0.748475 |

Equilibrium Haplotype frequencies

         B1       B2
A1 0.180975 0.538537 | 0.719512
A2 0.070549 0.209938 | 0.280487
   ------------------+---------
   0.251524 0.748475 |

Disequilibria, expectations and variances, haplotypes=200

D=-0.138158, ED=-0.137467, Var(D)=0.000245
EDmax=0.180070,Dmax=0.180975, Var(Dmax)=0.000301
D'=-0.763409, Var(D')=0.017386
```