Exam Project

Technical background information on how the data were generated

Genetics

2500 families consisting of

- 2 parents
- 0, 1, 2, 3 or 4 children, with 20% probability each

1000 diallelic genetic markers with

- Minor allele frequencies of 0.5 for markers A, B and C; 0.1 for markers D and E; 0.1 for markers F and F+1; and randomly taken between 0.01 and 0.5 for all other markers
- Hardy-Weinberg equilibrium for all markers
- No Linkage Disequilibrium between any pair of markers except between markers F and F+1
- Linkage Disequilibrium between markers F and F+1 modeled by defining correlation r (taken to be $\sqrt{0.8}$)
- Mendellian transmission of haplotypes for marker pair F, F+1 (without recombination) and of independent alleles for all other markers

Continuous phenotype

	Effects present	**Models**	**SNPS**
(A,B)	Pure epistasis, no main effects		(107,304)
(B,C)	Pure epistasis, no main effects	Different epistasis model than for (A,B)	(304,564)
(D,E)	Both interaction and main effects present	Same model as for (A,B)	(285,890)
F	Main effect	Additive model	645
F+1	Main effect	Advantage heterozygous model, SNP in LD with F	646

Assignment

Write an individual report which summarizes all your findings (homeworks 1-5) and discusses the results in the light of the foregoing information about the truth.

Compile your written report around the following items (not exclusively):

- Data quality control: e.g., pedigree structure, loops in data, missing data, Hardy-Weinberg equilibrium, kinships
- Main effects analysis:
    - Comparison population based and family based approach
    - Why are there more /less / other main effects seen?
    - Can the simulated models (additive, advantage heterozygous models) be retrieved? Why or why not?
    - What is the effect of LD between SNPs on main effects results?
    - What is the effect of the actual epistasis on main effects results?
- Interaction analysis:
    - Comparison of methods
    - Why are there more/less/ other pairs picked up?
    - Can one see differences in epistasis effects between the pairs found? Can you explain?
    - What is the effect of LD to discover interacting pairs?
    - What is the effect of an interacting pair belonging to a higher-order complex interaction system (A,B,C)?

The written report is handed in electronically before the exam date (as discussed in class)

Use the mail subject title "genetic epi exam project".

During the oral part we further discuss this report.