

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 $\sqrt{0.8}$)
- Mendelian 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