536 Homework 6

Adrian Dobra adobra@uw.edu

Table 1 cross-classifies the same individuals **4970** individuals based on their disease status and their genotypes at loci SNP 1 and SNP 2. We denote by X_1, X_2 and X_3 the three variables represented in this table, namely "Disease", "SNP 1" and "SNP 2". Both SNPs are biallelic: their minor alleles are denoted *a* and *b*, while their wildtype alleles are denoted *A* and *B*. Perform an analysis of this dataset using log-linear models and interpret what you learned from your analysis. Give an interpretation of each possible log-linear model in the context of the data from Table 1 (even for those models that do not fit the data well).

Disease (X_1)	SNP 1 (X_2)	*BB* (1)	*Bb* (2)	*bb* (3)
			SNP 2 (X_3)	
No (1)	*AA* (1)	1167	377	186
	Aa (2)	763	225	130
	aa (3)	107	29	11
Yes (2)	*AA* (1)	1509	16	179
	Aa (2)	234	2	19
	aa (3)	14	0	2

Table 1: An example of a $2 \times 3 \times 3$ table for genotype-disease association.