Biostatistics 666
Problem Set 5
Due November 27, 2017

Power of Genomewide Association Studies.

Assume that you are evaluating the power of potential genomewide association study where 1,000,000 independent genetic markers will be measured in 2,000 cases of type 2 diabetes and 2,000 population controls.

a) What might be an appropriate significance level for this study?

b) To help evaluate power, fill in the following matrix:

Power as a function of Allele Frequency and Effect Size

Population Allele Frequency (f)	Relative Risk		
	Modest (r=1.1)	Moderate (r=1.3)	Large (r=1.5)
Low (f = 0.05)	0%		
Moderate (f = 0.20)			
Common (f = 0.50)			

*Assuming a multiplicative model and assuming the population prevalence of type 2 diabetes is 10%.

c) What sample size would be required to achieve 80% power for a low frequency allele (f = 0.05) that makes a modest contribution to disease risk (r = 1.1)?

d) If your budget is limited and genotyping the number samples suggested in c) is outside your budget, how might you increase power for detecting alleles that make modest contributions to disease risk?

Affected Sibling Pair Analysis

In a study of type 2 diabetes (population prevalence = 5%), investigators collected 10 affected sibling pairs. These pairs were genotyped for a polymorphism with two alleles, allele "1" with frequency 0.20 and allele "2" with frequency 0.80. Genotyping results show that for all 10 pairs one sibling has genotype "2/2" and the other sibling has genotype "1/1".

a) If a genetic variant in the locus being examined had a population frequency of 10% and increased the risk of type 2 diabetes by 2-fold for heterozygotes and 4-fold for homozygotes, what would be the expected IBD proportions among affected sibling pairs?

b) Calculate the LOD score for this dataset using the MLS test of Risch (1990). Does the result suggest there is evidence for linkage?

c) Would the possible triangle constraint of Holmans (1993) affect the estimated LOD score? Why is the constraint useful?

d) Differences between reported and actual relationships among the individuals being studied can affect the power of a genetic linkage study. Speculate how this might be important.

e) Genotyping error is another common challenge for genetic studies. Speculate how this might be important.