Sample Size Calculation
Cochran-Armitage Test for Trend

Use the following web page to carry out the exercises:

http://ihg.helmholtz-muenchen.de/cgi-bin/hw/power2.pl

Question 1

You plan to collect 500 cases and 1,000 controls for disease X and wish to know whether or not this is a sufficient sample size to map disease susceptibility loci. Disease X has a population prevalence of 5%. You wish to estimate the power to detect a genotypic relative risk of 1.5 under a multiplicative model ($\gamma_2 = \gamma_1^2$) assuming a disease allele frequency of 0.08. What is the power for α=0.01? _____

Question 2

You wish to carry out a candidate gene study using 1,500 SNP marker loci and want to reject the null hypothesis using a p-value of 0.05. Determine what your power would be if you used a Bonferroni correction to control for the Family Wise Error Rate (FWER). Using the parameters provided in question 1 what is the power under the multiplicative model?

_____

Question 3

You determine that you are able to ascertain 750 cases and 750 controls what is the power using the same parameters as described in question 1? _____

Question 4

Like all statistical tests the power of the Cochran-Armitage test for trend is dependent on the underlying genetic model. Using the parameters from question 1 which of the following underlying genetic models: multiplicative ($\gamma_2 = \gamma_1^2$), additive($\gamma_2 = 2\gamma_1 - 1$), dominant ($\gamma_2 = \gamma_1$) or recessive ($\gamma_1 = 1$) would you predict to be the most powerful _____ and least powerful_____?

Question 5

Using the parameters described in question 1 except this time use a genotypic relative risk of 1.7, estimate the power under a multiplicative _____; additive _____; dominant_____; and recessive _____model.

Question 6

You have selected tagSNPs using a cut off of a minor allele frequency (MAF)=0.2. You estimated the allele frequency for your disease allele to be 0.08. If the disease and a SNP are in complete linkage disequilibrium (LD) what is the possible maximum r^2 between the two loci _____. Using the parameters in question 1 to have the same power how many cases _____and controls _____would you have to study. Hint: adjust the sample size using the maximum estimate of r^2.

Question 7

For a study design which uses equal number of cases and controls where tagSNPs have been selected using an $r^2 \geq 1$ and MAF ≥ 0.05 for a genotype relative risk of 1.8 under an additive for a disease with a population prevalence of 0.03 and disease allele frequency of 0.1. How many cases and controls should you ascertain for $\alpha=0.01$ and $1-\beta=0.80$? _____

ANSWERS

Question 1

What is the power for α=0.01? ___0.78___

Question 2

Using the parameters provided in question 1 what is the power under the multiplicative model? ____0.21____

Question 3

You determine that you are able to ascertain 750 cases and 750 controls what is the power using the same parameters as described in question 1? ____0.808_____

Question 4

Keeping all parameters consistent which of the following underlying genetic models: multiplicative ($\gamma_2 = \gamma_1^2$), additive($\gamma_2 = 2\gamma_1 - 1$), dominant ($\gamma_2 = \gamma_1$) or recessive ($\gamma_1 = 1$) would you predict to be the most powerful _multiplicative_ and least powerful__recessive____?

Question 5

Using the parameters described in question 1 except this time use a genotypic relative risk of 1.7, estimate the power under a multiplicative ___0.975_____; additive _____0.96_____; dominant ____0.929_____ and recessive _____0.015_____model.

Question 6

If the disease and a SNP are in perfect linkage disequilibrium (LD) what is the maximum r^2 between the two loci __0.348__. Using the parameters in question 1 to have the same power how many cases __1437____and controls __2874____would you have to study.

Question 7

How many cases and controls should you ascertain for an α=0.01 and 1-β=0.80 __319 cases + 319 controls = 638__?