Pardis Sadatian Moghaddam

Panther ID: 002722641

HW4:

1. **Consider loci 1 and 2. Suppose the allele for locus 1 is either A or a and the allele for locus 2 is either B or b. Suppose A and B are associated such that (1) pAB = pApB + D1, (2) pAb = pApb – D2, (3) paB = papB –D3, and (4) pab = papb +D4. Can you show that D1 =D2 =D3 =D4?**

    pAB = pApB + D1

    pAb = pApb – D2

    paB = papB – D3

    pab = papb + D4

    D1 = pAB – pApB
    P A = pAB + p Ab
    p B = pAB + p aB
    p a = pab + p aB
    pb = pab + p Ab

    D2 = pApb – pAb = pA(1-pB) – pAb = pA – pApB – pAb = pAB + pAb – pApB – pAb = pAB – pApB
    ⇨ D2 = pAB – pApB

So, we can say D1=D2

D3 = paB – papB = p aB – (1-pA)pB = paB – pB + pApB = p aB – [pAB + paB] + pApB = -pAB + pApB = -D1
⇨ D1 = pApB – pAB

So, we can say D1 = D3

D4 = papb – pab = (1-pA)(1-pB)-pab = 1-pB-pA+pApB-pab = pb – pA + pApB – pab = p ab + pAb – pA + pApB – pab = pAb – pA + pApB = pAb – pAB - pAb + PApB = -pAB + pApB
⇨ D1 = pApB – p(AB)

2. Consider the following two SNPs.
- SNP1: Case—40A,160C,Control—20A,180C,

- SNP2: Case—20A,180C,Control—10A,190C.

Can you compute the relative risk and the odds ratio for the two SNPs? Which SNP has a higher risk?

Case 1

	Case	Control
A	40 (p)	20 (q)
C	160 (r)	180 (s)

Relative risk => RR = [40/(160+180)]/[160/(40+20)] = 1.416

Odds ratio => OR = [40/160]/[20/180] = 1.37

RR > 1 and OR > 1 exposure is positively associated with disease.

Case 2

	Case	Control
A	20 (p)	10 (q)
C	180 (r)	190 (s)

Relative risk => RR = [p/(p+q)] / [r/(r+s)] = [20/(180 + 190)]/[180/(20+10)] = 1.37

Odds ratio => OR = [p/q]/[q/s] = [20/10]/[10/190] = 2.11

By comparing RR and OD in both cases, we can say case 1 has the higher risk.

2. Given a set of 6 haplotypes over 11 SNPs, find the minimum size subset of tag SNP which will distinguish all haplotypes
   the subset consists of SNPs:_____
   0,1,2,3,4,5,6,7,8,9,10
   (0,1,0,1,0,1,0,1,0,1,0)
   (0,1,1,0,1,0,1,0,0,0,0)
   (0,0,0,1,0,1,0,0,1,1,1)
   (1,0,1,1,0,1,0,1,1,0,1)
   (1,0,1,0,1,0,0,1,1,0,1)
   (1,0,0,0,1,0,1,0,1,1,1)
   ri(k) => as the allele of the i-th SNP in the k-th haplotype.

   We can partition 11 SNPs in to 3 blocks:

   r1,…,r3

   r4,…,r7

r8,…,r11

For r1,..,r3 is selected we should consider two things:
r1=0, we predict r1,..,r3 = 010 and 011 and 000 as the tag SNP; We predict the whole.
For r1=1, we predict r1,..,r3 = 101 as the tag SNP; This one can predict 3 out f 6 so it is more than 50%.

For r4,..,r7 is selected we should consider two things:
r1=0, we predict r4,..,r7 = 0101 as the tag SNP; We predict the whole with this tag.
For r1=1, we predict r4,..,r7 = 1010 as the tag SNP; This one can predict 5 out f 6 so it is more than 80%.

For r8,..,r11 is selected we should consider two things:
r1=0, we predict r8,..,r11 = 0111 as the tag SNP; We predict the whole with this tag.
For r1=1, we predict r4,..,r7 = 1101 as the tag SNP; This one can predict 5 out of 6 so it is more than 66%.