

Lösningar:

1. .

a. $\Omega = \{AB, Ab, aB, ab\}$

b. I populationen med n individer (kromosomer) finns m_i stycken av varje allel-kombination.

c. Låt (X, Y) vara den två-dimensionella s.v. som beskriver

försöket, där $X = \begin{cases} 1, & A \\ 0, & a \end{cases}, Y = \begin{cases} 1, & B \\ 0, & b \end{cases}$, om vi väljer en

individ från populationen, så är sannolikhetsfunktionen:

$$p_{XY}(x, y) = P(X = x, Y = y) = P(X = x)P(Y = y)$$

$$= p_A^x (1 - p_A)^{1-x} \cdot p_B^y (1 - p_B)^{1-y} \quad \text{om vi antar}$$

oberoende. (p_A och p_B motsvarar här den marginella allelfrekvensen för resp. locus)

d. Om

AB	Ab	aB	ab
0.4	0.25	0.2	0.15

Så blir korrelationen:

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = 0.4 - 0.65 \cdot 0.6 = 0.01$$

$$\rho = \frac{0.01}{\sqrt{0.65 \cdot 0.35 \cdot 0.6 \cdot 0.4}} = 0.0428$$

2. Låt A, B och C beteckna de tre populationerna, dvs.

$$P(A) = 0.2, P(B) = 0.5, P(C) = 0.3$$

$$P(S|A) = 4 \cdot 10^{-4}, P(S|C) = 0, P(S) = 10^{-4}$$

a. $P(A|A \cup B) = \frac{P(A)}{P(A) + P(B)} = \frac{0.2}{0.7} = 0.286$, eftersom A och B är disjunkta.

b. Med satsen om total sannolikhet får vi att

$$P(S|B) = \frac{P(S) - P(S|A)P(A)}{P(B)} = 4 \cdot 10^{-5} \text{ (dvs den relativa risken)}$$

blir 10 gånger större för någon i population A, jämfört med någon i population B)

c. Med Bayes Sats så följer att

$$P(B|S) = \frac{P(S|B)P(B)}{P(S)} = \frac{4 \cdot 10^{-5} \cdot 0.5}{10^{-4}} = 0.2$$

d. Med informationen från c. Så har vi att

$X = \text{antal sjuka från population B,}$

är $\text{Bin}(100, 0.2)$, med väntevärde 20, alltså förväntar vi oss 20 sjuka från population B och 80 från population A, men ingen från population C.

3. I en region på kromosom 4 är intensiteten för överkorsningar 0.9 cM/Mb (cM=1% chans att två loci har rekombinerats under en generation)

a. Intensiteten för en region på 100 bp blir

$$\frac{100bp}{10^6 bp} 0.9 \cdot 0.01 = 9 \cdot 10^{-7} \text{ överkorsningar/generation.}$$

b. X kan antas vara Poissonfördelad.

$$F_{Y_i}(y) = 1 - e^{-\lambda y}$$

c. sannolikheten att

$$F_{\min(Y_1, Y_2, Y_3, Y_4)}(z) = 1 - (e^{-\lambda z})^4 = 1 - e^{-4\lambda z}$$

avståndet mellan överkorsningarna är minst i den andra kromosomen blir 0.25, p.g.a. symmetri så har alla 4 lika stor sannolikhet att vara minst!

d. Använd normalapproximation, vilket ger:

$$P(\bar{X}_n < 0.01) = P\left(\frac{\bar{X}_n - \lambda}{\sqrt{\lambda/1000}} < \frac{0.01 - \lambda}{\sqrt{\lambda/1000}}\right) \approx \Phi(333) = 1$$

(Om vi skulle beräkna den sannolikheten för en större region skulle denna sannolikhet bli mindre. Pröva att göra om beräkningen om regionen istället var 10 kb (10 000 bp) lång.)