Statistical methods in genetic relatedness and pedigree analysis

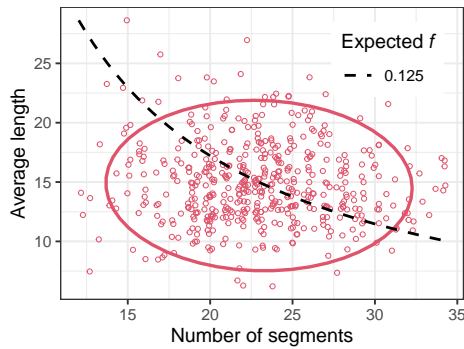NORBIS course, Oslo, June 2022
Magnus Dehli Vigeland and Thore Egeland

Solutions for exercise set IV

Note: Some answers are given in the exercise and therefore omitted here.

Exercise IV-1   (Realised inbreeding)

a) (Answer omitted.)

b) Haplotypes on chromosome 1 are shown below.

```
s = sims[[1]]
haploDraw(x, s, chrom = 1, cols = 2:7, pos = 2, height = 4, margin = c(2,1,1,1))
```


c) There are 2 autozygous segments on chromosome 1.

```
findPattern(s, pattern = list(autozygous = 7)) |> head()
```

```
##        chrom       start         end     length 7:p 7:m Aut
## [1,]       1   32.751940  38.574866   5.822926   4   4   1
## [2,]       1   70.691098  87.423440  16.732341   1   1   1
## [3,]       3    0.524278   3.516481   2.992203   4   4   1
## [4,]       3  116.151381 150.789104 34.637723   4   4   1
## [5,]       4    0.596033   4.265774   3.669741   3   3   1
## [6,]       4   72.690771  94.685256 21.994485   3   3   1
```
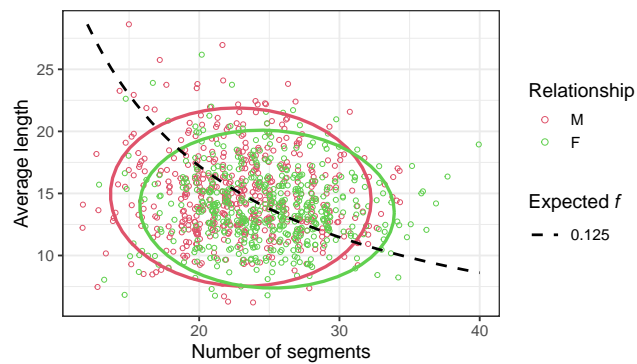
d) Typically 20-25 autozygous segments, but numbers outside that range are not uncommon:

```
plotSegmentDistribution(sims)
```

e)
```
x2 = swapSex(x, 4)
sims2 = ibdsim(x2, N = 500, ids = 7, seed = 321) # different seed
```

```
plotSegmentDistribution(M = sims, F = sims2, legendInside = F)
```


The distributions are different, with more/shorter segments in the version with most female meioses. However, the distributions are too overlapping to allow confident classification in any given case.

Exercise IV-2   (Realised IBD between siblings)

Complete R code:

```
x = nuclearPed(2)
sims = ibdsim(x, N = 500, ids = 3:4, seed = 123)
k = realisedKappa(sims)
showInTriangle(k$perSimulation)
```
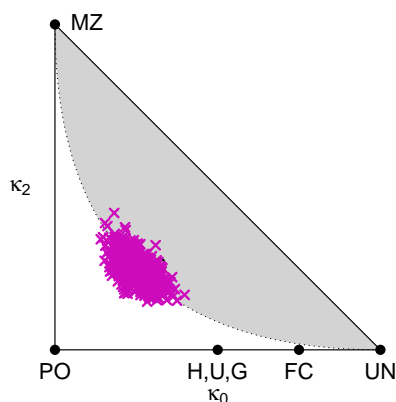
Exercise IV-3   (DNA sharing between siblings)

a) (Answer omitted.)

b) Possible R code:

```
phi = realisedKinship(sims)
sharedDNA = 2 * phi$perSimulation
mean(sharedDNA)
```

```
## [1] 0.4997271
```
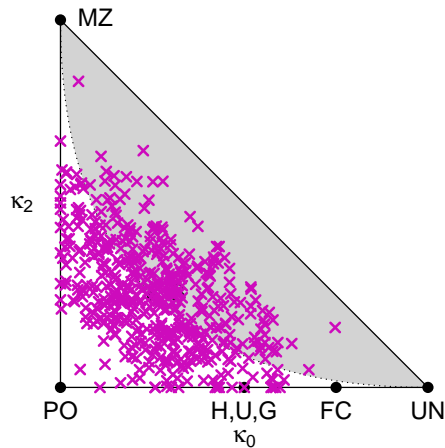
```
sd(sharedDNA)
```

```
## [1] 0.0406092
```

c) *The fraction of autosomal DNA shared IBD between siblings is 50% on average, with a standard deviation of ~4%.*

Exercise IV-4   (Fruit fly siblings)

Complete R code:

```
mapFF = list(uniformMap(cM = 107, chrom = 1),
             uniformMap(cM = 110, chrom = 2))
x = nuclearPed(2)
simsFF = ibdsim(x, N = 500, ids = 3:4, map = mapFF, seed = 123)
kFF = realisedKappa(simsFF)
```

```
showInTriangle(kFF$perSimulation)
```


Exercise IV-5   (Zero IBD)

a)
```
x = linearPed(10)
sims = ibdsim(x, N = 500, ids = c(1, 21))
zeroIBD(sims)
```

```
## $zeroprob
## [1] 0.64
##
## $stErr
## [1] 0.02146625
```

These simulations estimate a probability of 0.64 that none of Napoleon's DNA survived down to you. *Comment.* Your answers may differ slightly from mine because of the random number generation. To ensure reproducible results, you can set an explicit random number seed in the simulation command, e.g., `ibdsim(..., seed = 1234)`.

b) To simulate a female line, we add `sex = 2` to the `linearPed()` call. Note also that we use the female founder (2) instead of the male (1) in the `ids` argument.

```
x = linearPed(10, sex = 2)
sims2 = ibdsim(x, N = 500, ids = c(2, 21))
zeroIBD(sims2)
```

```
## $zeroprob
## [1] 0.474
##
## $stErr
## [1] 0.02233043
```

The simulations estimate a probability of 0.47 of no sharing in the female case. The probability is lower in this case because of the higher recombination rate in females.