

Pedigree analysis in R

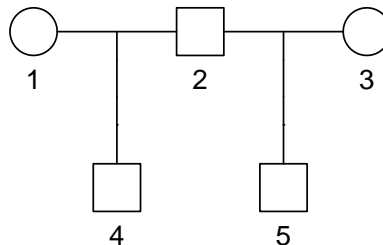
Magnus Dehli Vigeland and Thore Egeland

Solutions for exercise set IV

```
library(ribd)
library(ibdsim2)
```

Exercise IV-1

```
a) x = halfSibPed()
plot(x)
```

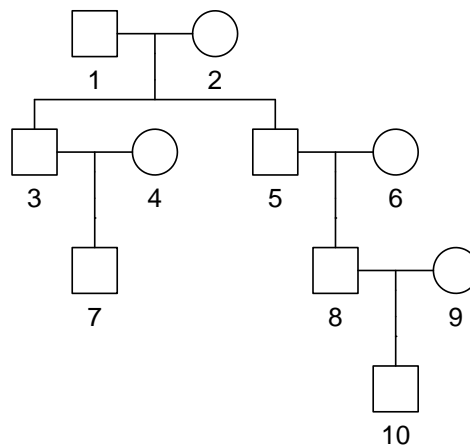


Kinship coefficient of the bottom individuals:

```
kinship(x, ids = 4:5)
```

```
## [1] 0.125
```

```
b) x = cousinPed(1, removal = 1)
plot(x)
```



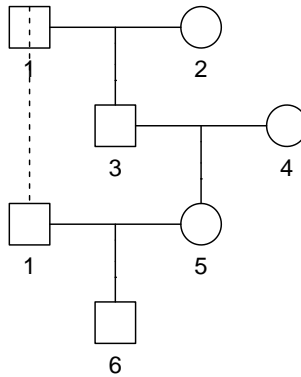
The kinship coefficient:

```
kinship(x, ids = c(7, 10))
```

```
## [1] 0.03125
```

Exercise IV-2

```
a) x = linearPed(2, sex = 1:2)
x = addChildren(x, fa = 1, mo = 5, nch = 1)
plot(x)
```



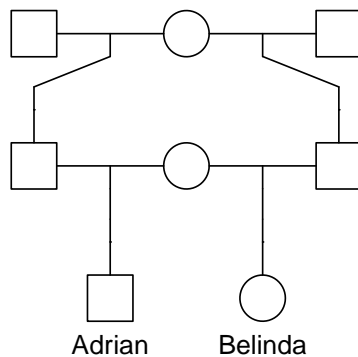
b) `inbreeding(x, ids = 6)`

```
## [1] 0.125
```

Exercise IV-3

a) Maternal half siblings, and half first cousins through the fathers.

```
b) x = halfSibStack(2)
   x = swapSex(x, 8)
   plot(x, labs = c(Adrian = 7, Belinda = 8))
```



c) $\phi = \frac{5}{32} = 0.15625$, as shown by this code:

```
kinship(x, ids = leaves(x))
```

```
## [1] 0.15625
```

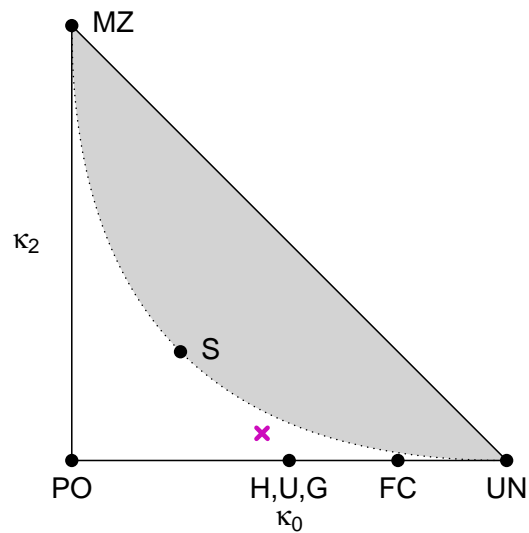
d) $\kappa = (\frac{7}{16}, \frac{1}{2}, \frac{1}{16})$, in agreement with the following calculation:

```
k = kappaIBD(x, ids = leaves(x))
```

k

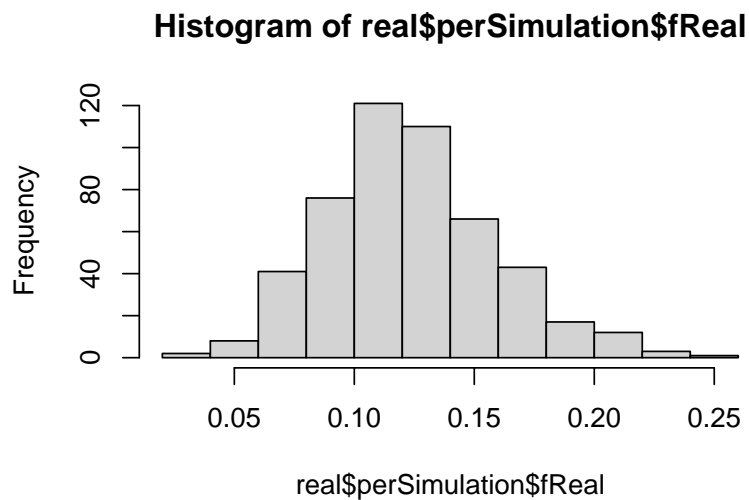
```
## [1] 0.4375 0.5000 0.0625
```

```
e) showInTriangle(k, lwd = 3)    # `lwd` adjusts symbol thickness
```



Exercise IV-4

```
a) x = cousinPed(deg = 0, removal = 1, child = TRUE)
sims = ibdsim(x, N = 500, ids = 7)
real = realisedInbreeding(sims)
hist(real$perSimulation$fReal)
```



The histogram shows substantial spread in the realised inbreeding coefficient in children of avuncular parents. The distribution seems fairly normal.

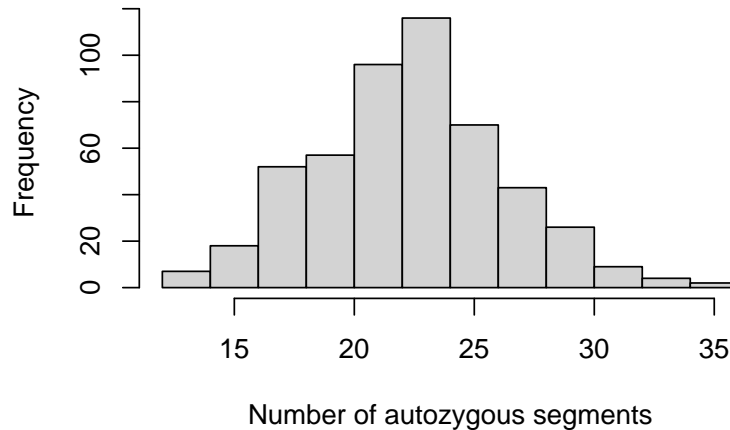
Note: Summary stats for the coefficients are easily obtained, for example with the command

```
summary(real$perSimulation$fReal)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.03856 0.09960 0.12031 0.12287 0.14298 0.24555
```

- b) The output table `real$perSimulation` has a column `nSeg` containing the number of autozygous segments in each simulation. We plot a histogram of this, including a more informative label on the x-axis:

```
hist(real$perSimulation$nSeg, xlab = "Number of autozygous segments", main = NULL)
```



Exercise IV-5

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

```
## $zeroprob
## [1] 0.628
##
## $stErr
## [1] 0.02161555
```

These simulations estimate a probability of 0.63 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] 0.448
##
## $stErr
## [1] 0.02223942
```

The simulations estimate a probability of 0.45 of no sharing in the female case. The probability is lower in this case because of the higher recombination rate in females.

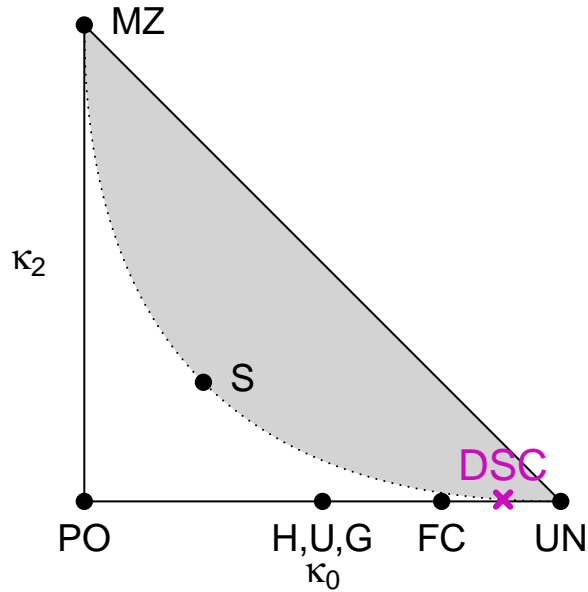
Exercise IV-6

a) R code:

```
x = doubleCousins(2, 2)
k = kappaIBD(x, ids = leaves(x))
k
```

```
## [1] 0.87890625 0.11718750 0.00390625
```

b) `showInTriangle(k, lwd = 3, label = "DSC", cexLab = 1.3, pos = 3)`



c) It is hard to tell from the plot if the point lies on the border or not. To prove that it does, we show that $\kappa_1^2 = 4\kappa_0\kappa_2$. Note that vector indexing in R starts at 1, hence `k[1]` is κ_0 as so on.

```
k[2]^2 == 4 * k[1] * k[3]
```

```
## [1] TRUE
```