Kinship and pedigree analysis: Methods and applications

Magnus Dehli Vigeland and Thore Egeland

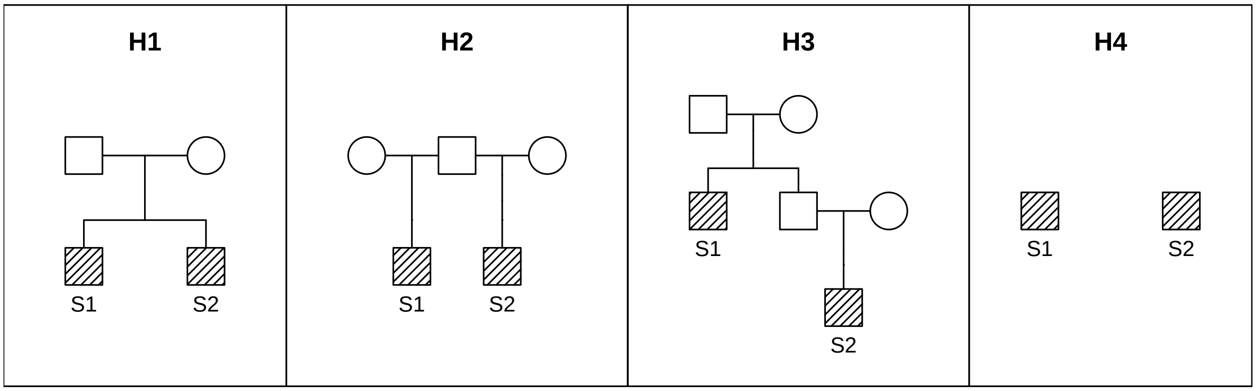
Exercise IV-1

The first part of this exercise is similar to Exercise III-2, but rather than the `pedsuite()`, we will be using `Familias`. In a paternity case, the alleged father (AF) has genotype a/b for a certain marker, while the child (CH, male) has genotype a/c. The allele frequencies are $p_a = 0.01$, $p_b = 0.3$ and $p_c = 0.69$ and the marker is called `M1`. The mother (MO) is not genotyped in the first part of the exercise

a) Perform step 1 of Familias: Enter the database.
b) Perform step 2 of Familias: Define the persons.
c) Perform step 3 of Familias: Enter the case data.
d) Perform step 4 of Familias: Define hypotheses H1 and H2 and calculate LR.
e) Assume the mother has genotype c/c. Calculate LR.
f) We consider a third hypothesis H3: A full brother, BR, of AF is the father. Find the LR comparing H1 to H3. *Hint*: Introduce persons GM (female) and GF (male) and define them as parents of both BR and AF. Define BR and MO as parents of CH.
g) Plot H3 using `QuickPed` to check that the pedigree was defined correctly.

Exercise IV-2

We revisit exercise III-3, but now using using `Familias`. Recall that we would like to investigate the relationship between two males, S1 and S2, who are genotyped with 15 STR markers. The hypothesised pedigrees are shown below.

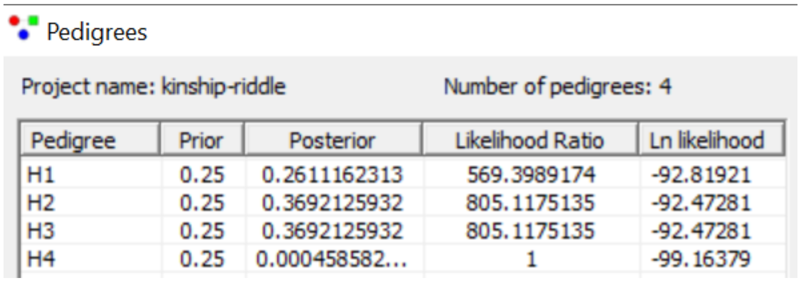


a) Load the file `kinship-riddle.fam` (available from the course home page) into Familias and verify the LRs from the previous exercise set , i.e.,

The posterior probabilities can be found using the following version of Bayes theorem

$$P(H_i \mid \text{data}) = \frac{\pi_i LR_{i4}}{\pi_1 LR_{14} + \pi_2 LR_{24} + \pi_3 LR_{34} + \pi_4}, \ i = 1, 2, 3, 4.$$

The prior probabilities for the hypotheses are $P(H_i) = \pi_i$ and LR_{i4} is the likelihood ratio comparing H_i to H_4.



Pedigrees

Project name: kinship-riddle

Number of pedigrees: 4

Pedigree	Prior	Posterior	Likelihood Ratio	Ln likelihood
H1	0.25	0.2611162313	569.3989174	-92.81921
H2	0.25	0.3692125932	805.1175135	-92.47281
H3	0.25	0.3692125932	805.1175135	-92.47281
H4	0.25	0.000458582...	1	-99.16379

Figure 1: Familias output in IV-1 a)

b) Verify the `Posterior` in the Familias output. Interpret the posterior $P(H_3 \mid \text{data})$.

c) We are informed that the age difference between S1 and S2 is 25 years. Based on this, we use the prior $\pi = (0.01, 0.09, 0.45, 0.45)$. Find $P(H_3 \mid \text{data})$ using Familias.

d) Include a hypothesis H5 specifying that S1 and S2 are first cousins, and find the LR against H4.

Exercise IV-3

This *optional* exercise demonstrates how you can continue projects in Windows `Familias` in R. Once the `Familias` data, i.e., the `.fam` file, has been converted to pedigree objects in R, we have access to all the functionality of the `ped suite`.

A child (CH) and the alleged father (AF) are genotyped with 21 STR markers. The mother is not disputed. We consider the following two hypotheses:

- H1: The alleged father is the biological father.
- H2: The alleged father and the child are unrelated.

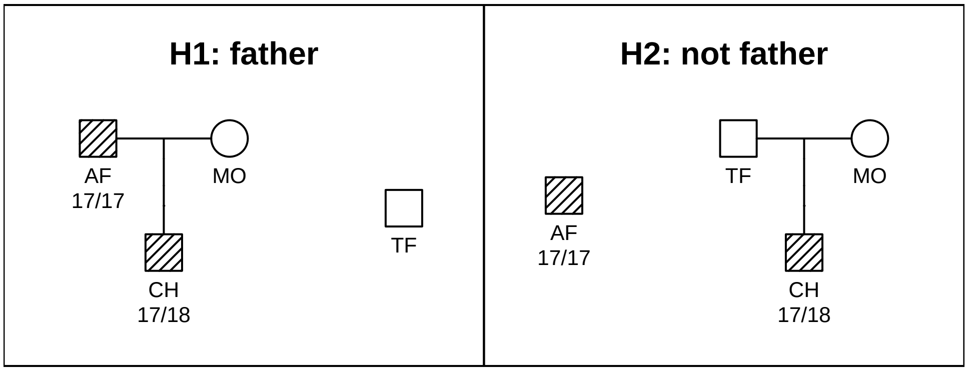
a) Use the `readFam()` function to download the Familias file and convert to R.

```
dat = readFam("http://familias.name/norbisRelatedness/paternityCase.fam")
```

b) The next step is to understand the output by `readFam()`. Run the following commands to explore the `dat` object.

```
is(dat)
names(dat)
summary(dat)
plotPedList(dat)
```

c) Produce the plot below as accurately as possible, including hatched symbols for the typed members, and genotypes for the first marker.



d) To simplify matters below, extract the two hypotheses as separate objects, `H1` and `H2`:

```
H1 = dat[[1]]
H2 = dat[[2]]
```

e) Show that LR comparing H1 to H2 is 0, by running:

```
res = kinshipLR(H1, H2)
res
```

f) Find the marker with LR = 0.

g) Calculate the LR once more, after first removing the marker giving LR = 0. (Note: This is a practice we advise strongly against!)

Rather than removing incompatible markers, we introduce a mutation model. The possible mutation models include `custom`, `equal`, `proportional`, `stepwise` and `onestep` and are described in the documentation of `pedmut::mutationModel()`. Different models can be used for females and males. Note that `custom` is completely general as you can define the mutation matrix. Below we use the `proportional` and `equal` models.

h) Apply a proportional model to PENTA_E (only) and recalculate LR:

```
H2 = setMutmod(H2, marker = "PENTA_E", model = "proportional", rate = 0.00001)
lr = kinshipLR(H1, H2, source = 2)
```

What is the total LR now? What is the LR for the marker PENTA_E?

i) What is the total LR when the `proportional` model with `rate = 0.00001` is used for *all* markers?

j) What is the total LR when the `equal` model with `rate = 0.00001` is used for *all* markers?

k) Run the below commands and explain what it does. Comment on the plot.

```
mutrate = 10^(-c(1:6))

LR = sapply(mutrate, function(r) {
  H2 = setMutmod(H2, model = "equal", rate = r)
  kinshipLR(H1, H2, source = 2)$LRtotal[1]
})

plot(mutrate, LR, type = "b", log = "xy", main = "LR with mutation model")
```