Statistical methods in genetic relatedness and pedigree analysis

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

Introduction

Before starting, load the core **pedsuite** packages.

```
library(pedsuite)
```

If you haven't downloaded the datasets (needed for the last exercise) already, do so by running

```
url = "https://magnusdv.github.io/pedinr/datasets/data.zip"
download.file(url, destfile = "data.zip")
unzip("data.zip")
```

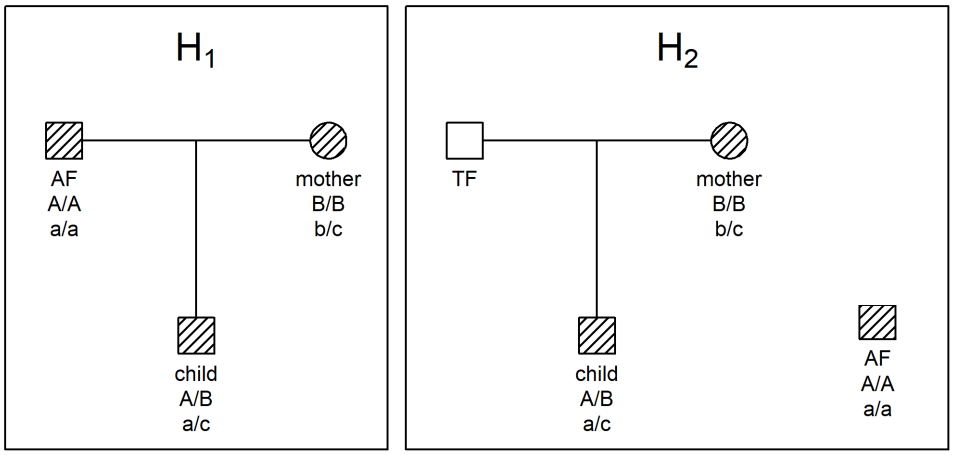
Exercise set V. Forensic exercises, part 1

Exercise V-1 (Paternity case)

The purpose of this exercise is to illustrate the basic concepts of forensic genetics and paternity testing by means of a simple paternity case. We consider the following two hypotheses

- H_1: The alleged father (AF) is the biological father.
- H_2: The alleged father and the child are unrelated.

The alleged father and the child are genotyped. The mother is not disputed. The figure below, with H_1 to the left, illustrates the hypotheses for the two markers used in this exercise.



a) Consider initially only the first marker. The allele frequencies are p_A and p_B. Explain why the likelihood ratio is $\text{LR}_1 = 1/p_A$.

b) Use R to plot LR_1 as a function of $p_A \in (0.001, 0.2)$. Explain why the likelihood ratio increases as p_A approaches 0.

c) There is a second autosomal marker, called S2, with alleles including a, b, and c with allele frequencies p_a, p_b, and p_c. Calculate the likelihood ratio for this marker and also for both markers.

d) Assume $p_A = 0.05$ and $p_a = 0.1$ and find the likelihood ratio in this case. How do you interpret LR?

e) Try the following R commands which can be used to check answers above:

```
# Pedigrees (hypotheses)
H1 = nuclearPed(fa = "AF", mo = "mother", ch = "child")
H2 = list(nuclearPed(fa = "TF", mo = "mother", ch = "child"),
          singleton("AF"))

# Marker 1
afr1 = c(A = 0.05, B = 0.05, C = 0.9)
H1 = addMarker(H1, AF = "A/A", mother = "B/B", child = "A/B", afreq = afr1)

# Marker 2
afr2 = c(a = 0.1, b = 0.1, c = 0.1, d = 0.7)
H1 = addMarker(H1, AF = "a/a", mother = "b/c", child = "a/c", afreq = afr2)

# Compute LR
kinshipLR(H1, H2, source = 1)
```

Exercise V-2 (Counterintuitive LR?)

We consider the same hypotheses as in the previous exercise. There is one marker with alleles denoted 1, 2, and 3. The alleged father is 1/2, the child is 1/3; the mother is untyped. The allele frequencies are p_1, p_2, and p_3.

a) Find the likelihood ratio.

b) Is it possible to have LR < 1 in a case where the father and son share an allele? Explain.

c) Use R to confirm the answers of the previous exercise when $p_1 = 0.5$, $p_2 = 0.3$ and $p_3 = 0.2$. A possible first part is:

```
# Pedigrees
H1 = nuclearPed(father = "AF", children = "CH")
H2 = list(singleton("AF"), singleton("CH"))

# Add marker to H1
H1 = addMarker(H1, AF = "1/2", CH = "1/3",
               afreq = c("1" = 0.5, "2" = 0.3, "3" = 0.2))

# You finish the code!
```

Exercise V-3 (Likelihood ratio for first cousins)

Consider the hypotheses

- H_1: A and B are first cousins.
- H_2: A and B are unrelated.

a) What is the LR if A and B share no alleles?

b) What is the smallest possible LR with 16 independent markers.

c) Consider an autosomal marker with alleles a, b, c and d with allele frequencies 0.1, 0.1, 0.1 and 0.7, respectively. What is the LR if both individuals are homozygous for allele d?

d) Repeat the previous calculation in R. A possible first part is:

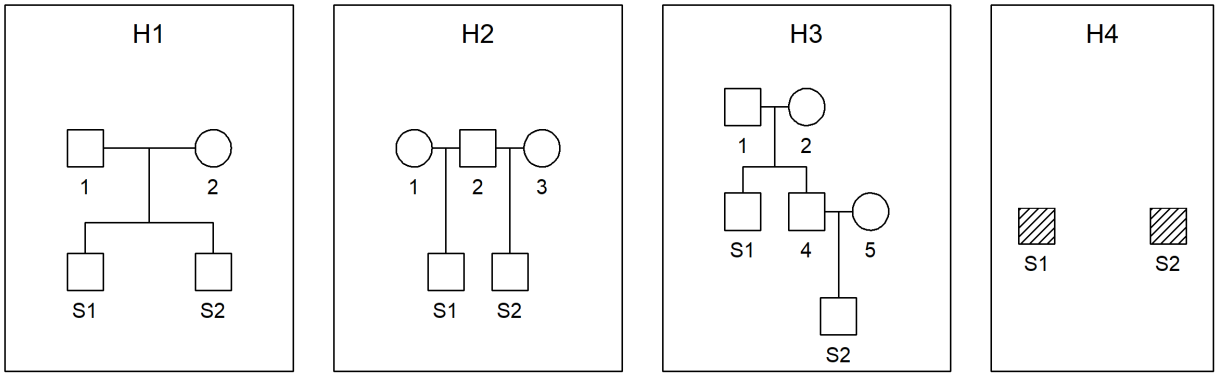
```
x = cousinPed(1) |>
  relabel(old = 7:8, new = c("A", "B"))

H1  = addMarker(x, A = "d/d", B = "d/d",
                afreq = c(a = 0.1, b = 0.1, c = 0.1, d = 0.7))
H2 = list(singleton("A"), singleton("B"))

# You finish the code!
```

Exercise V-4 (Relationship riddle)[1]

In this exercise we will conduct a test comparing four possible relationships between two male individuals, S1 and S2. The pedigrees are shown below.



a) Define the first three pedigrees by running:

```
ids = c("S1", "S2")
H1 = nuclearPed(children = ids)
H2 = relabel(halfSibPed(), old = 4:5, new = ids)
H3 = relabel(avuncularPed("uncle", "nephew"), old = c(3,6), new = ids)
```

b) The marker data for S1 and S2 is contained in the file *kniship-riddle.ped*, in the *data* folder. Load the data, and assign a frequency database, by running:

```
H4 = readPed("data/kinship-riddle.ped")
H4 = setFreqDatabase(H4, database = NorwegianFrequencies)
H4
```

c) Find the LRs when H4 is the reference.

d) Include a hypothesis specifying that S1 and S2 are first cousins and find the LRs. Use H4 as the reference. Comment on the result.
Hint:

```
H5 = relabel(cousinPed(1), old = 7:8, new = ids)
```

[1]Based on an example from the book "Pedigree Analysis in R", Vigeland, 2021, Academic Press, pp. 90 - 94.