Statistical methods in genetic relatedness and pedigree analysis

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

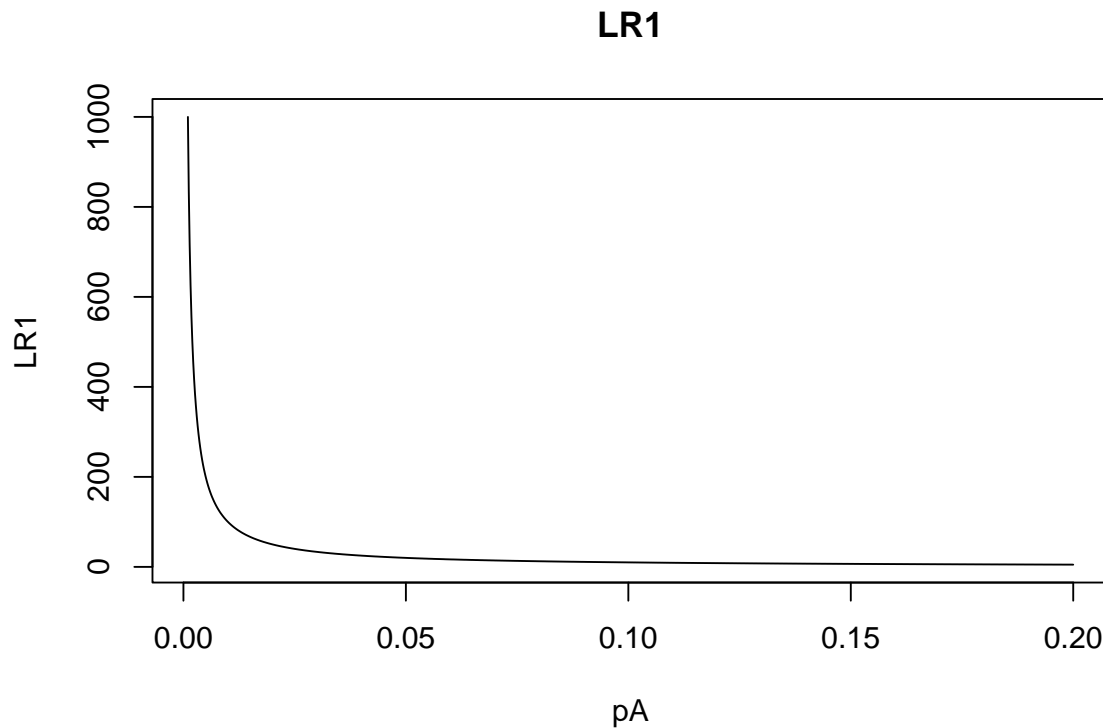## Solutions for exercises set V

Exercise V-1   (Paternity case)

a) We find

$$\text{LR}_1 = \frac{P(E \mid H_1)}{P(E \mid H_2)} = \frac{p_A^2 p_B^2 \cdot 1}{p_A^2 p_B^2 \cdot p_A} = \frac{1}{p_A}.$$

b) We can go

```
pA = seq(0.001, 0.2, length = 1000)
LR1 = 1/pA
plot(pA, LR1, type = "l", main = "LR1")
```


When the *paternal allele* A is rare, the evidence is strong as the alleged father and the child then share an uncommon allele; something unlikely to happen by chance.

c) We find

$$\mathrm{LR}_2 = \frac{P(E \mid H_1)}{P(E \mid H_2)} = \frac{p_a^2 2 p_b p_c \frac{1}{2}}{p_a^2 2 p_b p_c \frac{1}{2} p_a} = \frac{1}{p_a}.$$

The markers are independent and we may therefore multiply to get the combined likelihood ratio, i.e.,

$$\mathrm{LR} = \mathrm{LR}_1 \cdot \mathrm{LR}_2 = \frac{1}{p_A} \frac{1}{p_a} = \frac{1}{p_A p_a}.$$

d) In this case LR $= (1/0.05) \cdot (1/0.1) = 200$. The interpretation is that the data is 200 times more likely if we assume AF to be the biological father compared to the specified alternative.

e) (Answer omitted.)

Exercise V-2   (Counterintuitive LR?)

a) We find

$$\mathrm{LR} = \frac{2 p_1 p_2 \frac{1}{2} p_3}{2 p_1 p_2 \, 2 p_2 p_3} = \frac{1}{4 p_1}.$$

b) Yes, LR < 1 if $p_1 > 0.25$.

c) We can go

```
# 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))

# Compute LR
kinshipLR(H1, H2)

## H1:H2 H2:H2
##   0.5   1.0
```

Exercise V-3   (Likelihood ratio for first cousins)

a) We calculate the likelihoods using the expression

$$L(\kappa) = P(G_1, G_2 \mid UN)\kappa_0 + P(G_1, G_2 \mid PO)\kappa_1 + P(G_1, G_2 \mid MZ)\kappa_2.$$

For unrelated, $\kappa_0 = 1$, while for first cousins $\kappa_0 = \frac{3}{4}$, $\kappa_1 = \frac{1}{4}, \kappa_2 = 0$ as can be calculated using `ribd::kappaIBD`. Therefore

$$\mathrm{LR} = \frac{P(G_1, G_2 \mid UN)\frac{3}{4}}{P(G_1, G_2 \mid UN) \cdot 1} = \frac{3}{4}.$$

since $P(G_1, G_2 \mid PO) = P(G_1, G_2 \mid MZ) = 0$ when there is no allele sharing.

b) Smallest possible LR $= (3/4)^{16} \approx 0.01$ since markers are assumed to be independent.

c) If both are d/d, then

$$\mathrm{LR} = \frac{p_d^4 \frac{3}{4} + p_d^3 \frac{1}{4}}{p_d^4} = \frac{3}{4} + \frac{1}{4 p_d} = 31/28 = 1.10714.$$

d) We can go

```
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"))

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

```
##    H1:H2    H2:H2
## 1.107143 1.000000
```

Exercise V-4   (Relationship riddle)

a) (Answer omitted.)

b) (Answer omitted.)

c) We find

```
lr = kinshipLR(H1, H2, H3, H4,  source = 4, verbose = TRUE, ref = 4)
```

```
## Reference pedigree:  4
## Source pedigree:  4
## Number of markers:  15
```

```
lr
```

```
##    H1:H4    H2:H4    H3:H4    H4:H4
## 569.3989 805.1175 805.1175   1.0000
```

Comment: In this case only H4 has marker data attached. The argument `source = 4` secures that marker data from H4 are transferred to the other pedigrees. In this case it would work without `source = 4` since, as stated in the documentation, by default: "If exactly one of the pedigrees have attached markers, these are transferred to the others. All other cases give an error". I use `source = 4` for clarity, even if it is not needed.

d) We find

```
H5 = relabel(cousinPed(1), old = 7:8, new = c("S1", "S2"))
lr = kinshipLR(H1, H2, H3, H4, H5,  source = 4, verbose = TRUE, ref = 4)
```

```
## Reference pedigree:  4
## Source pedigree:  4
## Number of markers:  15
```

```
lr
```

```
##    H1:H4    H2:H4    H3:H4    H4:H4    H5:H4
## 569.3989 805.1175 805.1175   1.0000 110.0958
```