Likelihoods for pairs of markers

Thore

16 11 2019

Likelihoods

The likelihood of two individuals being related according to $\boldsymbol{\Delta}$, given their genotypes $G = (g_1, g_2)$ at a marker is obtained by conditioning on the Jacquard state:

$$L(\boldsymbol{\Delta} \mid G) = \sum_{i=1}^{9} \Delta_i P(G \mid J_i). \tag{1}$$

This likelihood is calculated efficiently for a large number of markers in `inbred::likJ`, function coded by Magnus.

The purpose of this section is to describe the extension to pair of linked markers. The further extension to independent pairs of such pairs of markers is trivial. Let $J^{(2)}$ denote the 9 by 9 matrix of identity states of a pair of pedigree members, for a given recombination rate. Furthermore, $\Delta^{(2)}$ is the 9*9 matrix of two-locus condensed identity coefficients, for a given recombination rate. Now G contains marker data for both loci. By conditionining, the likelihood may be written

$$L(\boldsymbol{\Delta}, \boldsymbol{\Delta}^{(2)} \mid G) = \sum_{s,t=1}^{9} \Delta_{s,t}^{(2)} P(G \mid J_{s,t}^{(2)}). \tag{2}$$

If we assume LE,

$$L(\boldsymbol{\Delta}, \boldsymbol{\Delta}^{(2)} \mid G) = \sum_{s,t=1}^{9} \Delta_{s,t}^{(2)} P(g_1 \mid J_s) P(g_2 \mid J_t). \tag{3}$$

This can be written on matrix form, convenient for implementation. To this end, we let

$$u = (P(g_1 \mid J_1), \ldots, P(g_1 \mid J_9))$$
$$v = (P(g_2 \mid J_1), \ldots, P(g_2 \mid J_9))$$

and write

$$L(\boldsymbol{\Delta}, \boldsymbol{\Delta}^{(2)} \mid G) = u\Delta^{(2)} v^T. \tag{4}$$

The numerical values in the below example is confirmed by FamLink http://famlink.se/f_index.html and Merlin. First, the numerator hypotheses is plotted

```r
library(inbred) # https://github.com/thoree/inbred
library(pedtools)
library(ribd)
p = c(0.4,  0.6)
a  = c(1, 1, 1, 1)
b  = c(2, 2, 1, 2)
cc = c(1, 2, 2, 2)
d  = c(1, 2, 1, 1)
pa = p[a]
pb = p[b]
pc = p[cc]
```

```
pd = p[d]
H1 = halfSibPed(1)
als = 1:length(p)
m = list()
for (i in 1:length(a))
   m[[i]] = marker(H1, afreq = p, alleles = als,
            "4" = c(a[i], b[i]), "5" = c(cc[i], d[i]) )
H1 = setMarkers(H1, m)
plot(H1,m, skip.empty.genotypes = TRUE, shaded = typedMembers(H1))
```
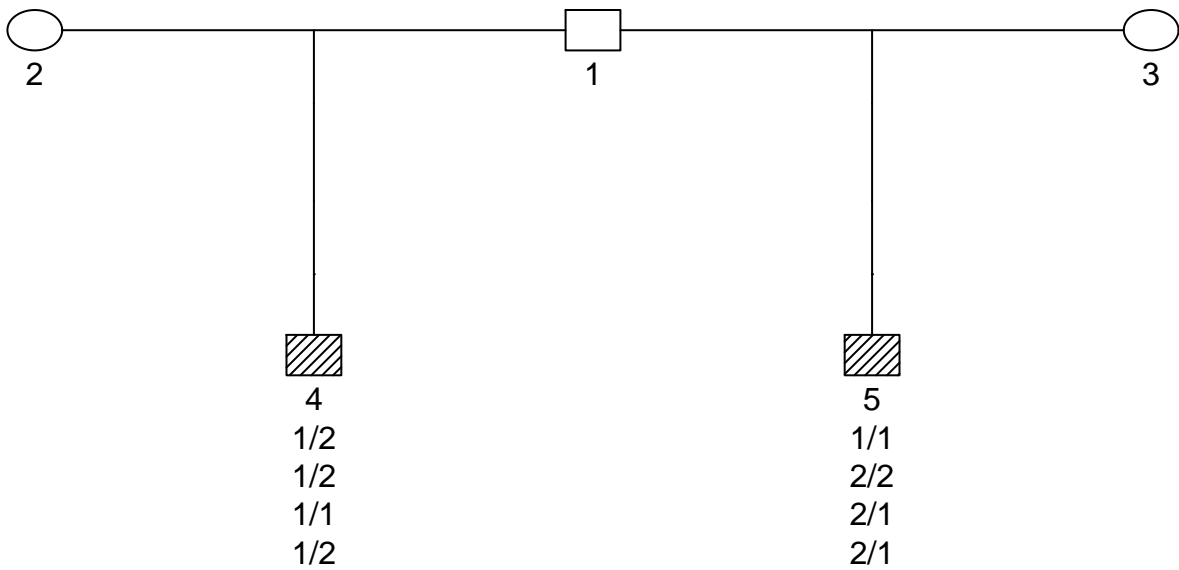

Next, the likelihood ratio is plotted as a function of the recombination rate in $[0, 0.1]$

```
Delta2 = matrix(0, ncol = 9, nrow = 9); Delta2[9,9] = 1
Delta1 = ribd::condensedIdentity(H1, c(4,5))
lik2 = likPairs(a,b,cc,d, pa, pb, pc, pd, Delta = Delta1, DeltaMatrix = Delta2)
denominator = prod(lik2[[2]])
rho = seq(0, 0.5, length = 100)
LRs = rep(NA, 100)
Delta1 = ribd::condensedIdentity(H1, c(1, 2))
for (i in 1:100){
  Delta2 = twoLocusIdentity(H1, c(4,5), rho[i])
  lik1 = likPairs(a,b,cc,d, pa, pb, pc, pd, Delta = Delta1, DeltaMatrix = Delta2)
  LRs[i] = prod(lik1[[2]])/denominator
}
plot(rho, LRs, type = "l", xlab = "rho", ylab = "LR")
```
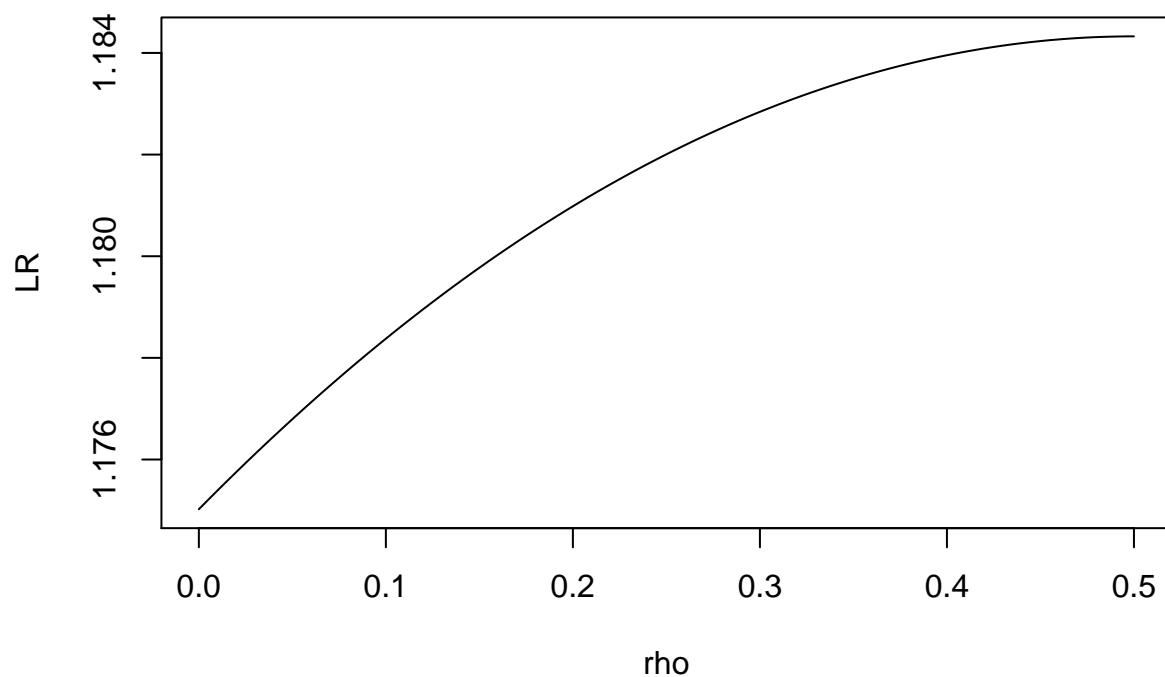
If the pedigree is defined in **pedtools** as above, the input can be extracted:

```r
ids = c(4,5)
nM = nMarkers(H1)
g = selectMarkers(H1, 1:nM)
odd = seq(1, nM*2, by = 2)
even = seq(2, nM*2,by = 2)
g2 = getAlleles(g)
a = g2[ids[1], odd]
b = g2[ids[1], even]
cc = g2[ids[2], odd]
d = g2[ids[2], even]
pa = p[as.integer(a)] # assumes integer alleles
pb = p[as.integer(b)]
pc = p[as.integer(cc)]
pd = p[as.integer(d)]
```