\[ P_{mn} = P_{m | n = m} \]

@ \( n = 1 \), can't have differences, so must split first. \( \Rightarrow P_{10} = 1 \) & \( P_{11} = 0 \)

\begin{align*}
\text{Mutation rate} \ @ \ & \frac{\Theta n}{2} \\
\text{Coal rate} \ @ \ & \frac{n(n-1)}{2}
\end{align*}

Competing rates so:

\[
P_{(\text{mutation})} = \frac{\Theta n}{2} = \frac{\Theta}{\Theta + n(n-1)} = \frac{\Theta}{\Theta + n-1}
\]

\[
P_{(\text{coal})} = \frac{n-1}{\Theta + n-1}
\]

\[
P_{mn} = P_{m | n = m} \text{ coal next } P_{(\text{coal})} + P_{m | n = m \text{ mut next}} P_{(\text{mut})} \text{ for } n = 2
\]

\[
P_{mn} = P_{n-1,m} \left( \frac{n-1}{n-1+\Theta} \right) + P_{n,m-1} \left( \frac{\Theta}{n-1+\Theta} \right)
\]

as req.

To find \( P_{mn} \) for a given \( n \),

start @ \( n = 1 \) & use the above equation

to increase \( n \) & \( m \) as required.

by summing over all paths to \( P_{mn} \)

& weighting by probability.
\[ S = 18 \quad k = 14 \quad n = 55 \quad \text{Tran} \text{. freq. column} \]

\[ \hat{\theta} = \frac{\sum_{i=1}^{j} \frac{1}{i}}{\frac{1}{2j}} = \frac{18}{4.5754} = 3.9346 \ldots \]

\[ 1 + \sum_{j=1}^{n-1} \frac{\theta}{\theta + j} = k = 14 \]

\[ \hat{\theta} = 5.73. \]

\( b \) set row \( k = 0 \) & all the entries are 0 if row, col = \( k, k \), else 1.

This example is in the notes as a binary matrix.

Gusfield's: order binary vectors from high to low, then read down paths to root.

from path to root: a 1 14 6 4
b 10 1 14 6 4
c 2 18

d 14 6 4

e 5 14 6 4
f 12 11 18

g 13 12 11 18
h 17 16 18
i 18
j 15 7
(a) 4-gene test: 4 7 (aka Corollary 5.6)

```
b 0 0
c 1 0
d 0 1
f 1 1
```

As this is present, infinite-sites not compatible.

ie. we cannot construct a tree.

eg. b is root, above already ordered

so paths to root are: b 4

<table>
<thead>
<tr>
<th></th>
<th>4</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

But there is no place to put f

so that it has mutation at 7 but not 4.

(or equivalently, place f but can't place c)

1 2 3 4 5 6 7

(b) a 0 0 1 1 1 0 0 0 (as binary)

b 0 0 0 0 0 0 0 0
c 0 0 1 1 0 0 0
d 1 1 1 1 0 1 1
 e 0 1 1 1 0 1 1
f 0 1 1 0 0 1 1

Now reorder:

```
3 4 5 2 6 7 1
```

a 1 1 1 0 0 0 0
b 0 0 0 0 0 0 0
 c 1 1 0 0 0 0 0
 d 1 1 0 1 1 1 1
 e 1 1 0 1 1 0
f 1 0 0 1 1 0 0

As diagram:

```
4
/
|
3
|
|
2
|
|
1

b
c
|   |
|---|---|
| a | 7 |   |
| b | 4 |   |
| c | 4 |   |
| d | 4 |   |
```
paths (from RHS):  a  5430
                   b  0
                   c  430
                   d 1762 430
                   e  762 430
                   f  762 30

Draw tree:

To place f, put additional mutation @ 4
back to ancestral type.
4. Sketch scenario first:

\[ \text{MRCA type} \]

\[ n-1 \text{ coals @ } k=3, \ldots, n \]

\[ m \text{ mutations @ } k=2 \]

\[ 1 \text{ final coal @ } k=2 \]

Finally, there are \( n \) choices for the MRCA.

\[ P(\text{sample}) = n \cdot P(A) \cdot P(B) \cdot P(C) \]

\[ = n \prod_{k=3}^{n} \left( \frac{k-1}{k} + \frac{\theta k}{2} \right) \left( \frac{\theta}{2(1+\theta)} \right)^m \left( \frac{1}{1+\theta} \right) \]

\[ = n \prod_{k=3}^{n} \frac{(k-1)(k-2)}{k(k-1+\theta)} \left( \frac{\theta}{2(1+\theta)} \right)^m \left( \frac{1}{1+\theta} \right) \]

\[ \frac{(n-1)(n-2)}{n(n-1+\theta)} \frac{(n-2)(n-3)}{(n-1)(n-2+\theta)} \frac{(n-3)(n-4)}{(n-2)(n-3+\theta)} \ldots \frac{3.2}{2(1+\theta)(1+\theta)} \frac{1}{(1+\theta)} \]
\[ = 2 (n-2)! \left( \frac{\theta}{2(1+\theta)} \right)^m \prod_{k=1}^{n-1} \frac{1}{k+\theta} \]

or for alternative approach use

\[ P(M_2 = m) \text{ from } \theta \mathcal{H}(c X_i) \text{ of sheet } 1: \]

\[ p(M_2 = m) = \left( \frac{\theta}{1+\theta} \right) \left( \frac{1}{\theta+1} \right) \]

\[ \text{Prob. (all mutations on same lineage)} = \frac{2}{2^m} \]

\[ P(M_k = 0) = \prod_{k=3}^{n} \frac{k-1}{\theta+k-1} = \frac{n-1}{\prod_{k=3}^{n}} \frac{k}{k+\theta} \]

\[ . \quad \text{p (sample)} = \frac{2}{n-1} \left( \frac{\theta}{2(1+\theta)} \right)^m \prod_{k=1}^{n-1} \frac{1}{k+\theta} \]

\[ = 2 (n-2)! \left( \frac{\theta}{2(1+\theta)} \right)^m \prod_{k=1}^{n-1} \frac{1}{k+\theta} \quad \text{(as before)} \]