

Red = known

## Unknowns:

Probability that a person has Pneumonia, Tuberculosis, Lung Cancer and Bronchitis.

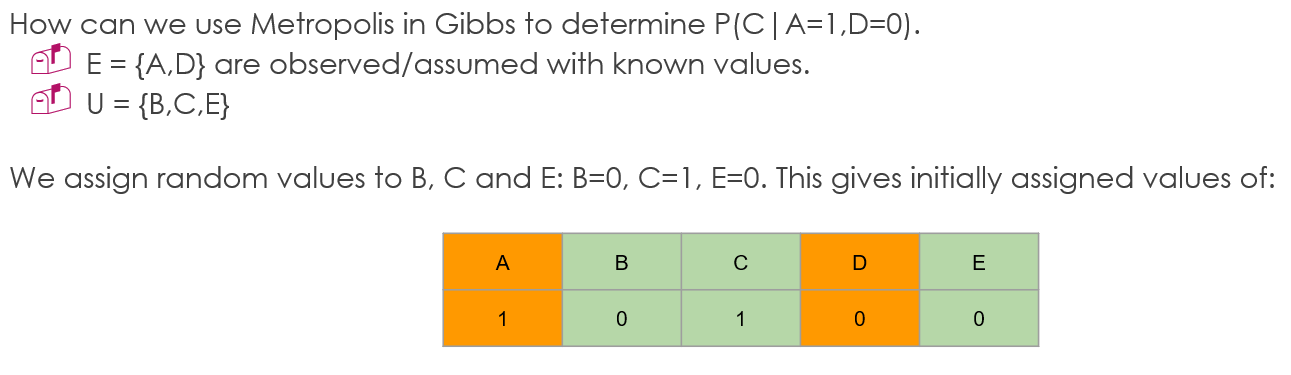
# Solution

### Create distribution table

From the file “historical data” create the distribution table for each node.

In the node temperature, the distribution table will contain mean and standard deviation for the cases Pneumonia = 0 and Pneumonia =1.

## Sampling



Our **observed** with known values: Te, VTB, Sm, XR, Dy, Our **Unknowns:** Pn, TB, LC, Br this ones we need to initialize with some value

Assigned values:

assign random values for Pn, TB, LC, Br

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Pn | Te | VTB | TB | Sm | LC | Br | XR | Dy |
| 1 | 38.73447 | 0 | 0 | 1 | 1 | 0 | 1 | 0 |

P\_old = P(Pn =1, Te=38.73447, VTB=0, TB=0, Sm=1, LC=1, Br=0, XR=1, Dy=0)

= **P(Pn =1)P(Te = 38.73447|Pn =1)P(VTB=0)P(TB=0|VTB=0)P(Sm =1)P(LC=1|Sm=1)P(Br=0|Sm=1) P(XR=1|Pn=1,TB=0,LC=1) P(Dy=0|LC=1,Br=0)** = 0,03 (example)

The temperature:

P(Te = 38.73447|Pn =1) = dnorm(38.73447, mean = 39, sd = 2)

Dnorm is the probability density function

|  |  |  |
| --- | --- | --- |
|  | Pn = 0 | Pn =1 |
| Te, mean | 37 | 39 |
| Te, standard diviation | 2 | 2 |

Example table

Proposed values:

We change the value Pn

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Pn | Te | VTB | TB | Sm | LC | Br | XR | Dy |
| 0 | 38.73447 | 0 | 0 | 1 | 1 | 0 | 1 | 0 |

P\_new= P(Pn =0, Te=38.73447, VTB=0, TB=0, Sm=1, LC=1, Br=0, XR=1, Dy=0)

= **P(Pn =0)P(Te = 38.73447|Pn =1)P(VTB=0)P(TB=0|VTB=0)P(Sm =1)P(LC=1|Sm=1)P(Br=0|Sm=1) P(XR=1|Pn=1,TB=0,LC=1) P(Dy=0|LC=1,Br=0)** = 0.02(example)

Choose P\_new or P\_old?

Since pnew < pold we accept the new value with probability pnew/pold= 0.02/0.03 = 2/3

We generate a (uniformly) random number between 0 and 1, and if it is less than 2/3 we accept the new value and assign it to Pn.

Then continue through all the unknown variable. That is one sample. Repeat for 5-10000 samples.

Burn in: throw away the first 10% or so samples, because of bias of the inital random values.