Præsentation

Anders

28/10/2020

We have the dataset ChestSim1000

```
data(chestSim1000, package="gRbase")
head(chestSim1000)
```

```
##   asia tub smoke lung bronc either xray dysp
## 1   no  no    no   no   yes     no   no  yes
## 2   no  no   yes   no   yes     no   no  yes
## 3   no  no   yes   no    no     no   no   no
## 4   no  no    no   no    no     no   no   no
## 5   no  no   yes   no   yes     no   no  yes
## 6   no  no   yes  yes   yes    yes  yes  yes
```

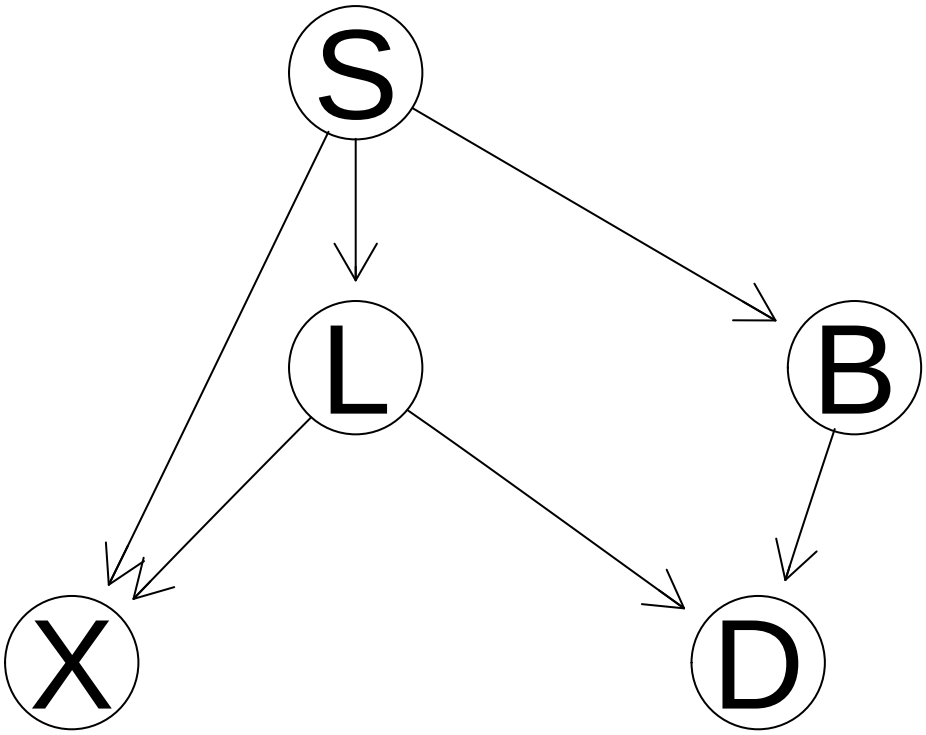
This is a hyphotetical Chest Clinic problem, by Lauritzen and Spiegelhalter. (ref til https://arxiv.org/pdf/1301.7394.pdf)

Here is a short explanation of the variables in the dataset.

- asia → subject has visited asia
- tub → subject has tuberculosis
- smoke → subject is a smoker
- lung → subject has lung cancer
- bronc → subject has bronchitis
- either → subject has either tuberculosis or lungcancer
- xray → subject has positive X-ray
- dysp → Subject has dyspnoea

Shortness-of-breath (dyspnoea) may be due to tuberculosis, lung cancer, bronchitis, none of them, or more than one of them. A recent visit to Asia increases the chances of tuberculosis, while smoking is known to be a risk factor for both lung cancer and bronchitis. The results of a single chest X-ray do not discriminate between lung cancer and tuberculosis, as does neither the presence nor absence of dyspnoea. (citat direkte sat ind fra https://arxiv.org/pdf/1301.7394.pdf)

```
dg1 <- dag(~ S + L|S + X|L:S + B|S + D|L:B)
plot(dg1)
```




```
P1 <- function(obj, i, j){
  x <- unlist(obj[which(names(obj) == i)])
  y <- unlist(obj[which(names(obj) == j)])
  a11 <- length(which(x == "yes" & y == "yes"))/length(x)
  a12 <- length(which(x == "no" & y == "yes"))/length(x)
  a21 <- length(which(x == "yes" & y == "no"))/length(x)
  a22 <- length(which(x == "no" & y == "no"))/length(x)
  mat <- matrix(c(a11,a12, a21,a22), nrow = 2, byrow = TRUE, dimnames = list(c("yes","no"),c("yes","no")
  names(dimnames(mat)) <- c(j,i)
  print(mat)
}
P1(chestSim1000, "asia", "tub")
```

```
##      asia
## tub     yes    no
##   yes 0.000 0.007
##   no  0.014 0.979
```

```
P1(chestSim1000, "smoke", "lung")
```

```
##      smoke
## lung    yes    no
##   yes 0.046 0.007
##   no  0.419 0.528
```