

Causal Graphs for Basic Epidemiologic Data

Part 5—Front-door criterion

Tomás J. Aragón

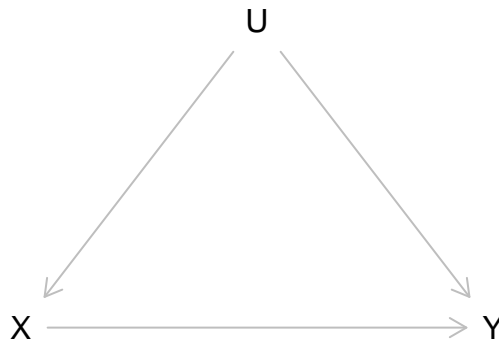
2016-11-28

Lung cancer case study

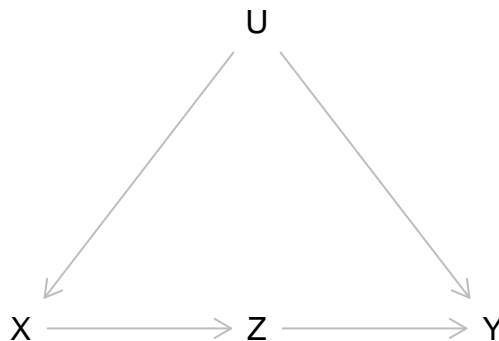
See p. 66 in Pearl book.

Case fully described here: <http://bayes.cs.ucla.edu/BOOK-2K/ch3-3.pdf>

```
library(dagitty)
g1 <- dagitty('dag{X->Y U->X U->Y}')
coordinates(g1) <- list(x = c(X = 0, U = 1, Y = 2),
                        y = c(X = 1, U = 0, Y = 1))
plot(g1)
```



```
g2 <- dagitty('dag{X->Z Z->Y U->X U->Y}')
coordinates(g2) <- list(x = c(X = 0, Z = 1, U = 1, Y = 2),
                        y = c(X = 1, Z = 1, U = 0, Y = 1))
plot(g2)
```



```
#### Create data frame
X.lv <- c("Smokers", "Non.smokers")
Z.lv <- c("Tar", "No.tar")
Y.lv <- c("Cancer", "No.cancer")
farray.yzx <- array(c(323, 57, 18, 2, 1, 19, 38, 342), dim = c(2, 2, 2),
                    dimnames = list(Y = Y.lv, Z = Z.lv, X = X.lv))
```

```

fdat <- epitools::expand.table(farray.yzx)
str(fdat)

## 'data.frame': 800 obs. of 3 variables:
## $ Y: Factor w/ 2 levels "Cancer","No.cancer": 1 1 1 1 1 1 1 1 1 1 ...
## $ Z: Factor w/ 2 levels "Tar","No.tar": 1 1 1 1 1 1 1 1 1 1 ...
## $ X: Factor w/ 2 levels "Smokers","Non.smokers": 1 1 1 1 1 1 1 1 1 1 ...

library(bnlearn)
dag.fd <- empty.graph(nodes = c("X", "Z", "Y"))
dag.fd <- set.arc(dag.fd, from = "X", to = "Z")
dag.fd <- set.arc(dag.fd, from = "Z", to = "Y")

bn.mle <- bn.fit(dag.fd, data = fdat, method = "mle")
bn.mle

##
## Bayesian network parameters
##
## Parameters of node X (multinomial distribution)
##
## Conditional probability table:
##      Smokers Non.smokers
##      0.5      0.5
##
## Parameters of node Z (multinomial distribution)
##
## Conditional probability table:
##
##      X
## Z      Smokers Non.smokers
## Tar      0.95      0.05
## No.tar    0.05      0.95
##
## Parameters of node Y (multinomial distribution)
##
## Conditional probability table:
##
##      Z
## Y      Tar No.tar
## Cancer  0.81  0.14
## No.cancer 0.19  0.86

(pY_Z.X <- prop.table(xtabs(~Y+Z+X, data=fdat),c(2,3)))

## , , X = Smokers
##
##      Z
## Y      Tar No.tar
## Cancer  0.85  0.90
## No.cancer 0.15  0.10
##
## , , X = Non.smokers
##
##      Z

```

```

## Y           Tar No.tar
##   Cancer    0.05  0.10
##   No.cancer 0.95  0.90

(pX <- bn.mle$X$prob)

##           Smokers Non.smokers
##           0.5      0.5

(pZ_X <- bn.mle$Z$prob)

##           X
## Z           Smokers Non.smokers
##   Tar        0.95      0.05
##   No.tar     0.05      0.95

#### Calculate front-door adjustment
#### Y = cancer among nonsmokers
pY1_do.X0.vec <- rep(NA, length(X.lv)*length(Z.lv))
y1 <- Y.lv[1] # cancer
x0 <- X.lv[2] # nonsmoker
step <- 1
for(xp in X.lv){
  for(z in Z.lv){
    pY1_do.X0.vec[step] <- pY_Z.X[y1, z, xp] * pX[xp] * pZ_X[z, x0]
    step <- step + 1
  }
}
(pY1_do.X0 <- sum(pY1_do.X0.vec))

## [1] 0.4975

#### Y = cancer among smokers
pY1_do.X1.vec <- rep(NA, length(X.lv)*length(Z.lv))
y1 <- Y.lv[1] # cancer
x1 <- X.lv[1] # smoker
step <- 1
for(xp in X.lv){
  for(z in Z.lv){
    pY1_do.X1.vec[step] <- pY_Z.X[y1, z, xp] * pX[xp] * pZ_X[z, x1]
    step <- step + 1
  }
}
(pY1_do.X1 <- sum(pY1_do.X1.vec))

## [1] 0.4525

```