DAGの例

$$\begin{tikzcd} D \arrow[rr] & & Y \\ & X \arrow[ul] \arrow[ur] & \end{tikzcd}$$

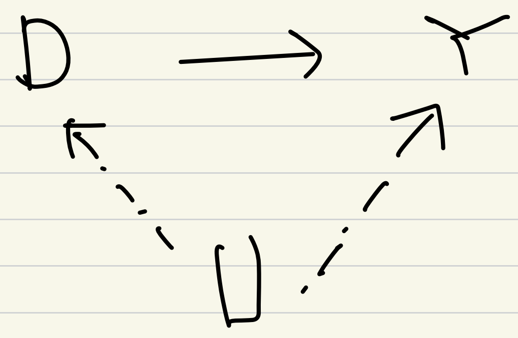
非合流点

$D \to Y$

$D \leftarrow X \to Y$

Dから Yを知りたいが、$D \leftarrow X \to Y$があると、Xの変動が、DもYも変動させる

$\longrightarrow$ DとYに疑似相関が生じる



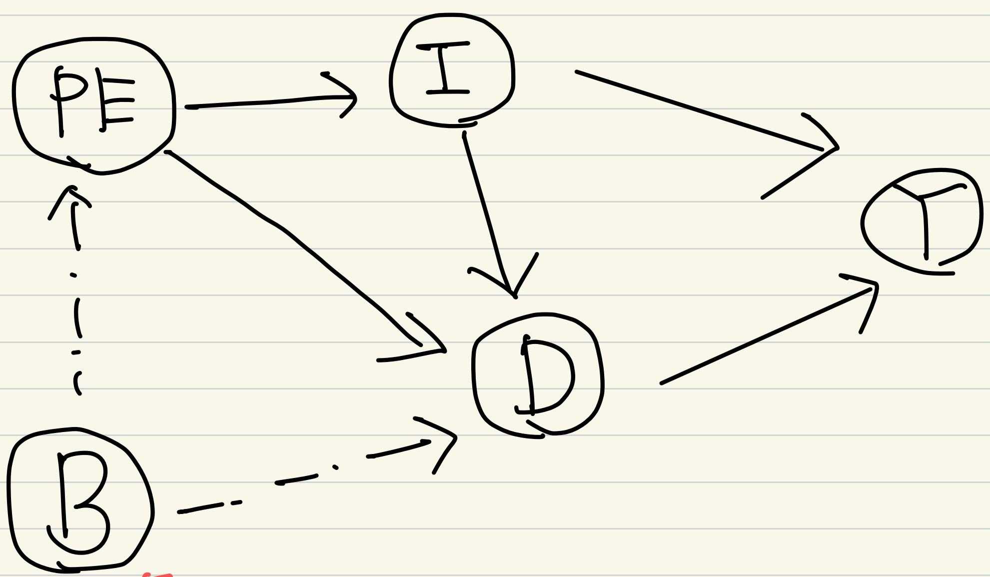
non collider.

未観測

$D \to Y$

$D \leftarrow U \to Y$

未観測のUがバックドアにいるとき.
バックドアがあいてるとどう.



未観測

知りたい効果

$D \to Y$

$D \leftarrow I \to Y$

$D \leftarrow PE \to I \to Y$.

$D \leftarrow B \to PE \to I \to Y$.

バックドア (合流点はないので全て開いてる)

↳ $D \to Y$の直接効果を推定できない.

$$\begin{array}{ccc} D & \longrightarrow & Y \\ & \searrow \quad \swarrow & \\ & X & \end{array}$$

collider

○ $D \to Y$

$D \to X \leftarrow Y$

バックドア

パスに沿って

2つの変数が第3の変数に影響する場合、
その変数を合流点と呼ぶ

バックパス上に合流点があることは、
バックパスは閉じることになる。

逆に条件付けすることで、逆にパスが開いてくる

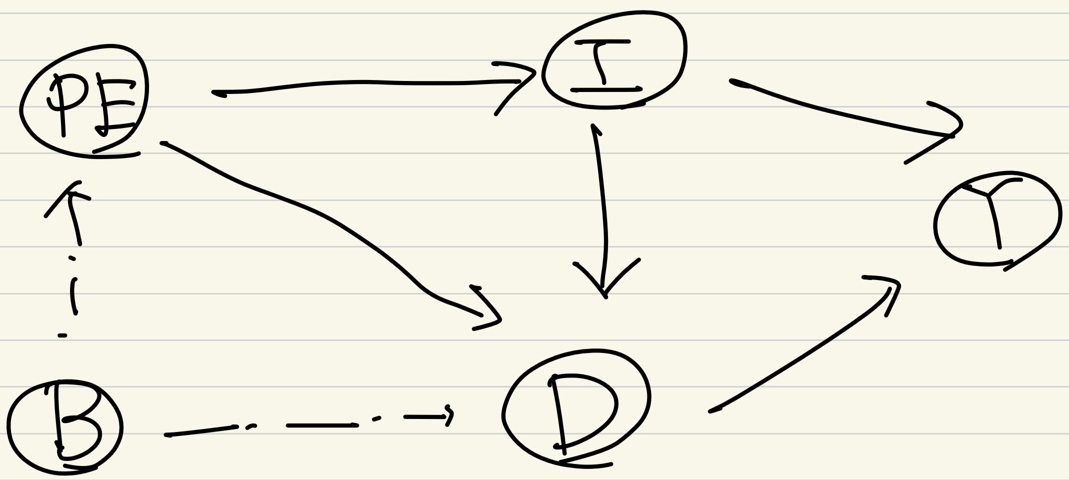
合流点は無視した方がよ…

バックパスを閉じるには?

① 未観測の因子を、マッチングなどを工夫して、無視出来るようにするため。

② バックドアパス上の合流点を作り無視することで バックパスを閉じる

全てのバックパスが閉じられたとき ⇒ バックドア基準を満たす。



D → Y.

D ←…B → PE → I → Y.

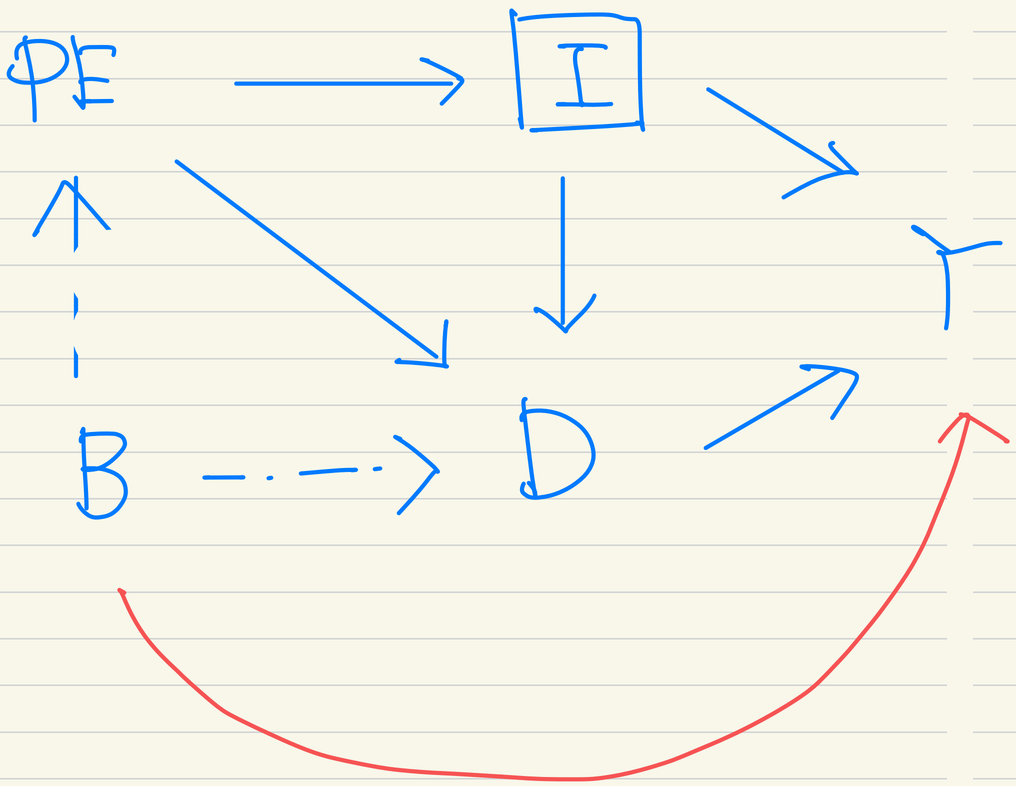
D ← PE → I → Y

D ← I → Y

Iが全てのバックドアの接点なので Iでの条件付けで全て解決。

$Y_i = \beta_0 + \beta_1 D + \beta_2 I_i + \epsilon$ でモデル化出来る。

一方で次のDAGを想定すると

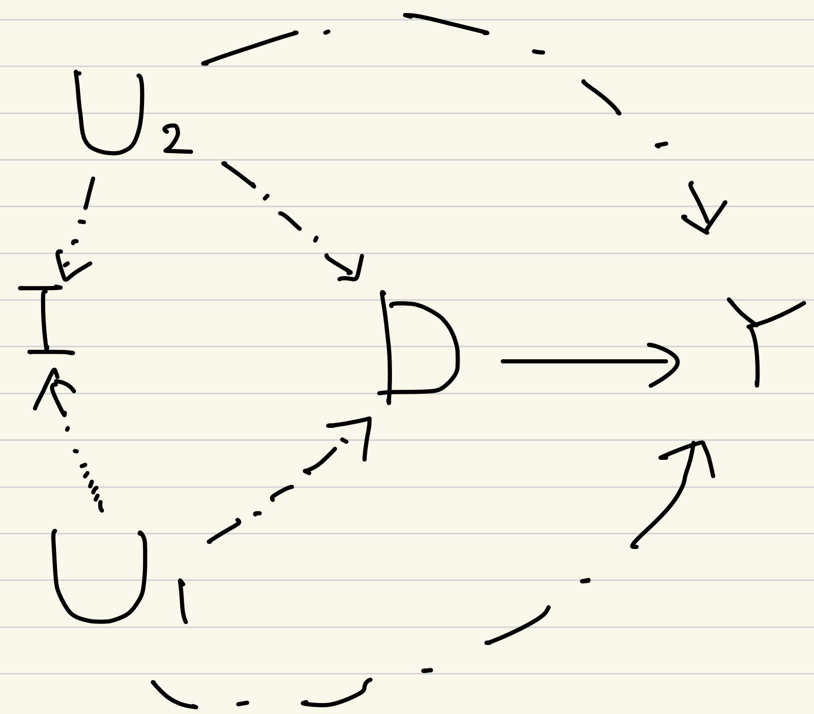


新たに、

D ←… B …→ Y というバックドアが

出来き，モデル化が出来なくなる

合流点バイアスの例



$D \rightarrow Y$

① $D \leftarrow U_2 \rightarrow Y$.

② $D \leftarrow U_1 \rightarrow Y$.

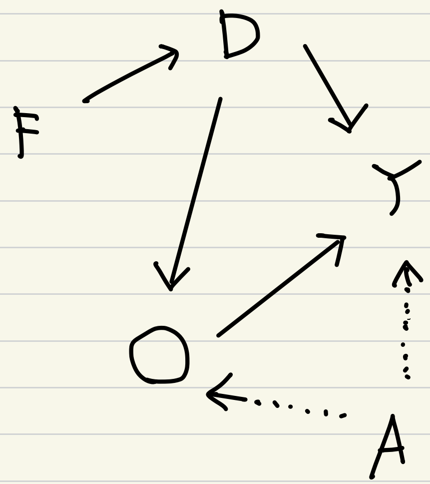
③ $D \leftarrow U_2 \rightarrow I \leftarrow U_1 \rightarrow Y$.

④ $D \leftarrow U_1 \rightarrow I \leftarrow U_2 \rightarrow Y$.

①,②,は未観測なので,DAGでは処理出来ない

③,④は合流点があるので閉じている ⟶ Iで条件付けて開いてしまう

⇒ このDAGだと,条件付けでは,バックドア基準をみたすことが出来ない.



$D \to Y$

① $D \to \bigcirc \to Y$

② $D \to \bigcirc \leftarrow A \to Y$

① … バックドアパスではない

② … ○が合流点となっており閉じている

→ ○で条件付けないと、D→Yの効果はわからないが、

○で条件付けると、観測できないAによる交絡が生じる。

→