- Table Representation

model a data with 8 binary features

↓ enumerate all possible combinations

| $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ | $X_8$ | $P$ |
|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | |
| | | | | | | | | |

problems:

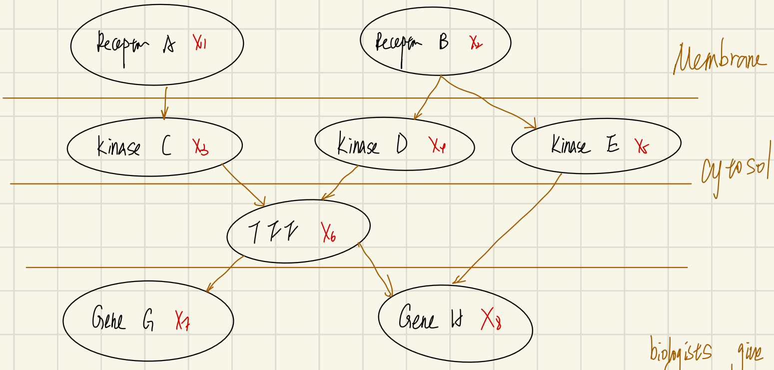
storing the table

filling the table

compute any quantity out of this table

- Multivariate Distribution in High-dim space

eg: Cellular signal transduction




biologists give you a causal model

combine domain-knowledge into a mathematical framework : Graphical Model

nodes: random variables

edge: relationship (how to define)

- Pearson's Correlation

Normalized covariance $\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$

captures linear dependency

linear regression from X to Y gives $\beta = \frac{Cov(X,Y)}{Var(X)}$

$X \perp\!\!\!\perp Y \rightarrow \rho(X,Y) = 0$

$\rho(X,Y) = 0 \not\rightarrow X \perp\!\!\!\perp Y$ eg $X \sim uniform(-1,1)$ $Y = X^2$ (non-linear dependencies)

$X, Y$ has 0 Pearson Correlation

Cannot capture non-linear dependency

one way to construct measure of dependence

$P(X,Y) \xleftrightarrow{\text{distance}} P(X)P(Y)$

distance $=0$ iff $X \perp\!\!\!\perp Y$

· KL - divergence

$$KL(P,Q) = \int_x p(x) \log \frac{p(x)}{q(x)} dx$$

mutual information: $I(X,Y) = KL(P_{XY}, P_X P_Y)$

$I(X,Y) = 0$ iff $X \perp\!\!\!\perp Y$

Hard to compute :(

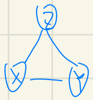
eg.

X = height of kid

Y = vocabulary of kid

Z = age of kid

$\xrightarrow{\text{Marginal correlation}}$




· Partial Correlation

partial correlation given a random vector Z

correlation measured between X and Y after eliminating linear effect of Z

i.e. correlation between residual from regression from Z to X and Z to Y

$$\rho(X,Y|Z) = \rho(e_X, e_Y) = \frac{Cov(e_X, e_Y)}{\sqrt{Var(e_X)}\sqrt{Var(e_Y)}}$$

$$e_X = X - (W_X^T Z + b_X)$$

$$e_Y = Y - (W_Y^T Z + b_Y)$$

(Z is the cause)

$X \perp\!\!\!\perp Y \mid Z \rightarrow \rho(X,Y|Z) = 0$

$\rho(X,Y|Z) = 0 \not\rightarrow X \perp\!\!\!\perp Y \mid Z$

$R_{ij} = \rho(X_i, X_j \mid X_k \cdots)$

if each variable is marginally a Gaussian,

then $P(X_1, \cdots, X_n) \sim N(\mu, \Sigma)$

the partial correlation matrix R can be computed through inverse of covariance matrix

$$R_{ij} = -\frac{\theta_{ij}}{\sqrt{\theta_{ii}}\sqrt{\theta_{ij}}}$$

Measure of association between X and Y

- Marginal
  - Linear
    - pearson's Correlation
  - Non-linear
    - $dist(P_{XY}, P_X P_Y)$
      - kL-divergence
        - mutual information
      - Max Mean discrepancy
        - Hilbert-Schmidt Independence Criterion
- Non-marginal
  - Linear
    - partial correlation
  - Non-linear
    - conditional independence

PGM = multivariate statistics + structure

GM = multivariate objective function + structure