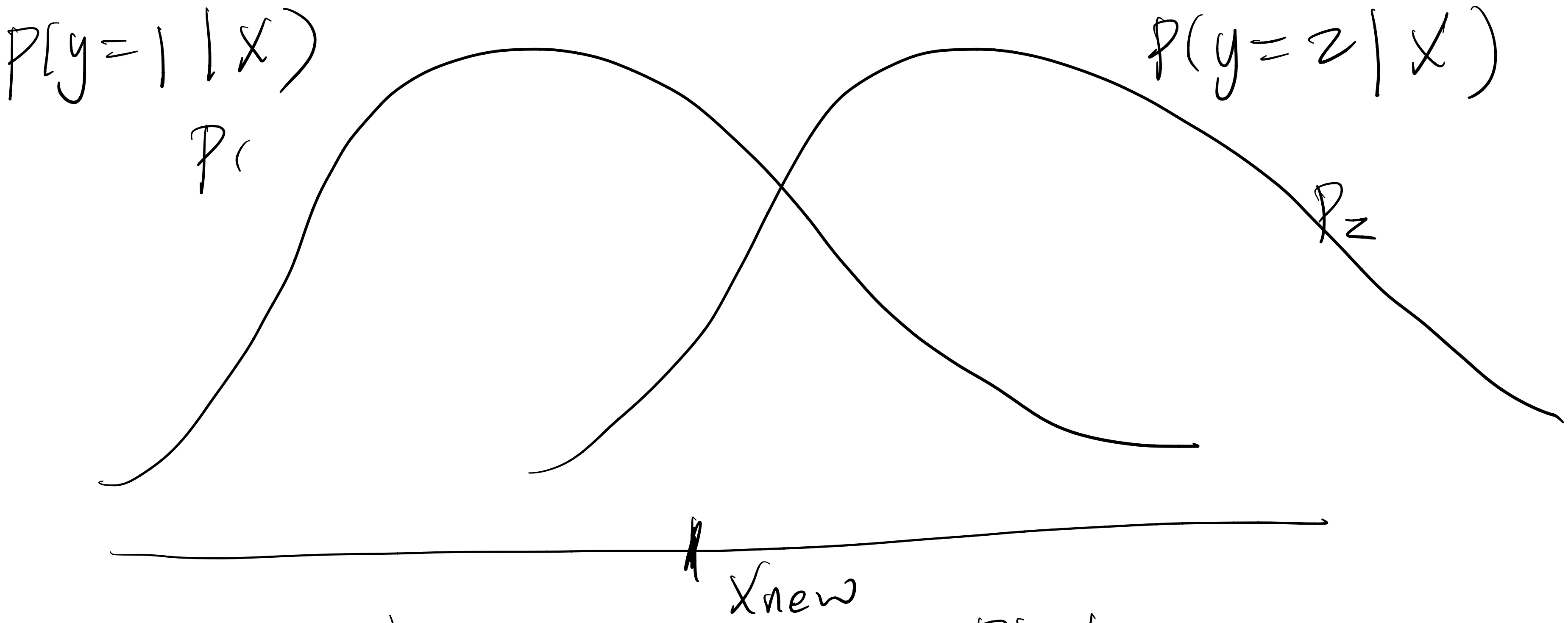
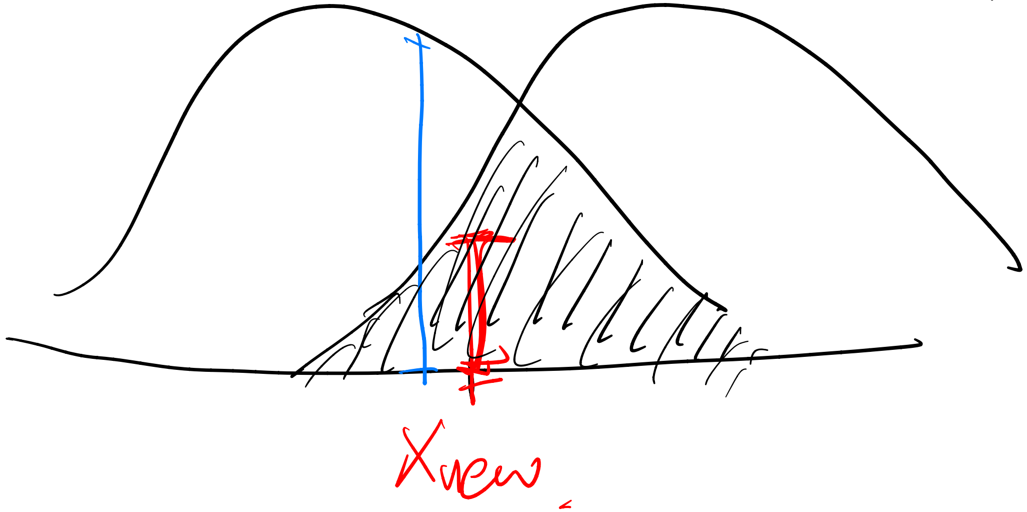
Bayes optimal classifier.



$$P_1(X_{new}) \gtrless P_2(X_{new})$$



Bayes error

$$\frac{P_2(X_{new})}{P_1(X_{new}) + P_2(X_{new})}$$

Risk $\quad R^* = \underbrace{\int \min\left[\frac{P_1(X)}{P_1(X)+P_2(X)}, \frac{P_2(X)}{P_1(X)+P_2(X)}\right]}\, P(X)\, dx$

if we assume $P_1(X) = P_2(X)$

$$P(X) = \frac{P_1(X) + P_2(X)}{2}$$

$$R^* = \frac{1}{2}\int \min\left[P_1(X), P_2(X)\right] dx$$

Graphical Models.

$$p(x) = \frac{1}{\sqrt{2\pi}^D \cdot |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$$

parameters : μ . Σ . $x \in \mathbb{R}^P$

$\mu \in \mathbb{R}^D$
$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix}$$

$\Sigma \in \mathbb{R}^{D \times D}$
$$\begin{bmatrix} \sigma_{11} & & \sigma_{1p} \\ & \sigma_{22} & \\ & & \ddots \\ & & & \sigma_{nn} \end{bmatrix}$$

$D +$

$$1 + 2 + \cdots + D = \boxed{\frac{D(D+1)}{2}}$$

If all elements x_i are independent each other?

$$\mu \in \mathbb{R} \quad \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_D \end{pmatrix} \qquad \Sigma = \begin{bmatrix} \sigma_{11} & 0 & 0 & \cdots \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & \sigma_{nn} \end{bmatrix} \rightarrow P(D) = 2D$$

If X_1 and X_2 are independent!

$$p(X) = p(X_1, X_2) = p(X_1) \cdot p(X_2)$$

$$\mu \in \mathbb{R} \qquad \Sigma = \begin{bmatrix} D_1 & \\ & D_2 \end{bmatrix} \qquad \frac{D_1(D_1+1) + D_2(D_2+1)}{2}$$

Fundamental Idea : Chain Rule.

$$p(x) = p(x_1, x_2, \cdots, x_D)$$

$$= p(x_1)\, \underbrace{p(x_2, x_3, \cdots, x_D \mid x_1)}$$

$$= p(x_1)\, p(x_2 \mid x_1)\, p(x_3, x_4, \cdots, x_D \mid x_2, x_1).$$

$$\vdots$$

$$= p(x_1)\, p(x_2 \mid x_1)\, p(x_3 \mid x_1, x_2)\, p(x_4 \cdots x_D \mid x_1, x_2, x_3)$$

If $x_1, x_2, \cdots, x_D$ are mutually independent.

$$p(X) \overset{①}{=} p(x_1)\, p(x_2 \mid x_1)\, p(x_3 \mid x_1, x_2) \cdots p(x_D \mid x_1 \cdots x_{D-1})$$

(as x_1 and x_2 are independent.

$$p(x_1, x_2) = p(x_1)\, p(x_2)$$

$$p(x_2 \mid x_1) = \frac{p(x_1, x_2)}{p(x_1)} = p(x_2)\ )$$

$$\overset{②}{=} p(x_1)\, p(x_2)\, p(x_3) \cdots p(x_D)$$

Directed Graphical Model (= Bayes Network).

"If a causality graph is given"

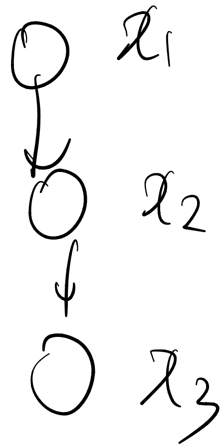
$$\begin{tikzcd}
x_1 \arrow[dr] & & x_3 \arrow[dl] \arrow[dr] & & & \\
& x_2 & & x_4, x_5 \arrow[dr] & & x_6 \arrow[dl] \\
& & & & x_7 &
\end{tikzcd}$$

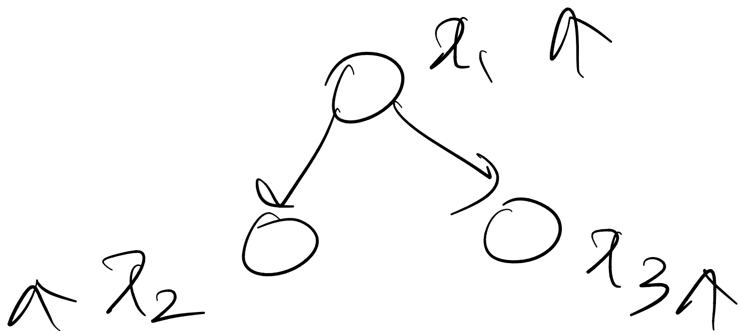
$$p(X) = p(x_1, x_2, \cdots, x_7)$$

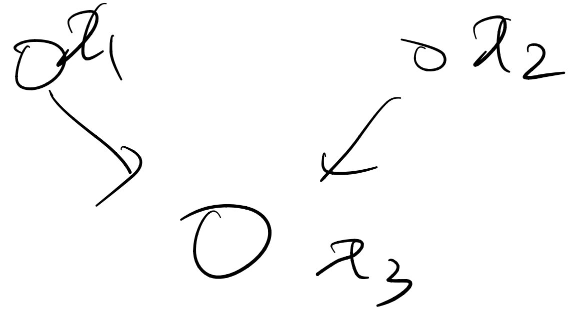
$$= p(x_1)\, p(x_3)\, p(x_2 | x_1, x_3)$$

$$\cdot p(x_4, x_5 | x_3)\, p(x_6)\, p(x_7 | x_4, x_5, x_6)$$

"Independencies" (D-separation)



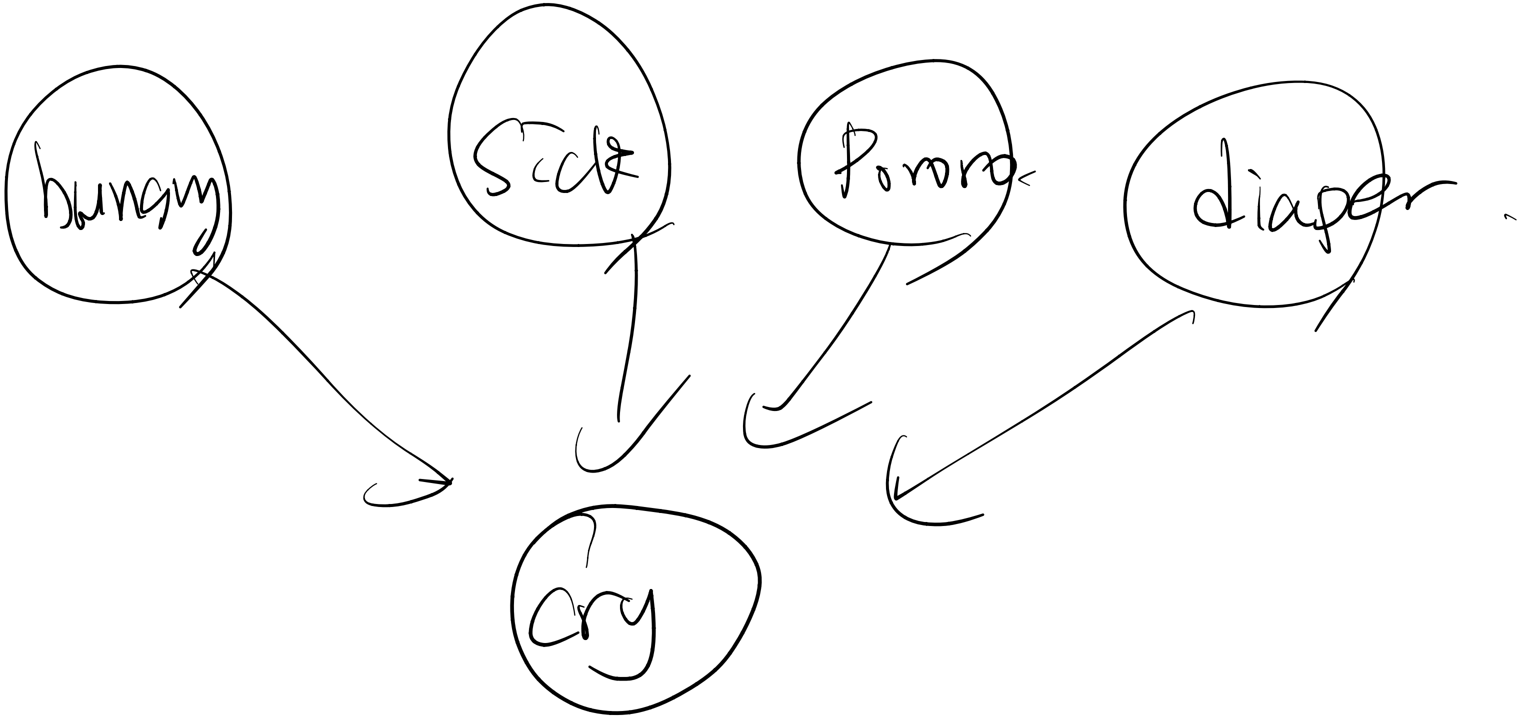




$x_1 \perp\!\!\!\perp x_3 \mid x_2$

$x_2 \perp\!\!\!\perp x_3 \mid x_1$

$$\begin{cases} x_1 \perp\!\!\!\perp x_2 \\ x_1 \not\perp\!\!\!\perp x_2 \mid x_3 \end{cases}$$


hungry
Sick
Pororo
diaper
cry