

Convolution

$$(\mu * \nu)(A) = \int \int_X \mathbb{1}_{\{x+y \in A\}} d\mu(x) d\nu(y)$$

$$= \mathbb{E}_{\mu \otimes \nu} [\mathbb{1}_{\{x+y \in A\}}]$$

λ Lebesgue

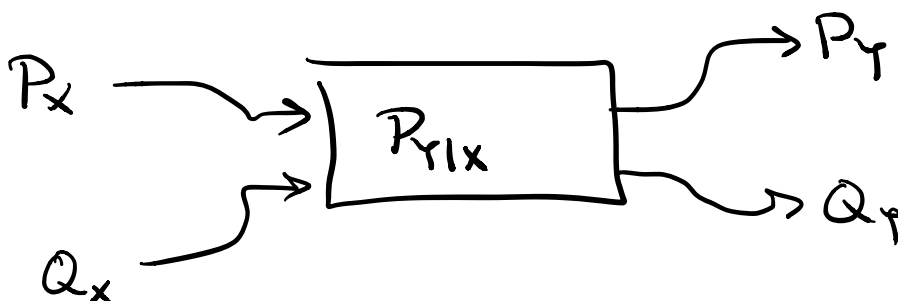
If μ and ν have densities $f = \frac{d\mu}{d\lambda}$ and $g = \frac{d\nu}{d\lambda}$:

$$(\mu * \nu)(A) = \int \int_X \mathbb{1}_{\{x+y \in A\}} f(x) g(y) dx dy$$

$$x+y=s \Rightarrow \int_X \mathbb{1}_{\{s \in A\}} \left(\int_X f(s-y) g(y) dy \right)$$

$$= \int_A (f * g)(s) ds$$

Data Processing Inequality



$$D_f(P_x \| Q_x) \geq D_f(P_Y \| Q_Y)$$

Item (i)

$$D_f(P_X \parallel Q_X) = D_f(P_{X,Y} \parallel Q_{X,Y})$$

$\uparrow \qquad \qquad \uparrow$
 $P_X P_{Y|X} \quad Q_X P_{Y|X}$

Projections

$$\text{Let } X = X_1 \times X_2$$

$$\text{Let } X = (X_1, X_2) \sim P_{X_1, X_2} \in \mathcal{P}(X_1 \times X_2)$$

$$X' = (X'_1, X'_2) \sim P_{X'_1, X'_2} \in \mathcal{P}(X_1 \times X_2)$$

$$\text{Let } Y = X_1 \quad \text{and} \quad P_{Y|X_1, X_2} = \delta_X,$$

namely,

$$Y = \text{proj}_1(X_1, X_2) = X_1$$

$$\text{DPI} \Rightarrow D_f(P_{X_1, X_2} \parallel Q_{X_1, X_2}) \geq D_f(P_{X_1} \parallel Q_{X_1})$$

$$\text{Item (iv)} \Rightarrow D_f(P_{X_1, X_2} \parallel Q_{X_1, X_2}) = D_f(P_{X_1, X_2, Y} \parallel Q_{X_1, X_2, Y})$$

$$\text{Note: } P_{X_1, X_2} = P_{X_1} P_{X_2|X_1} \quad ; \quad Q_{X_1, X_2} = Q_{X_1} Q_{X_2|X_1}$$

$$\text{If } Q_{X_2|X_1} = P_{X_2|X_1}, \text{ then}$$

$$D_f(P_{X_1, X_2} \parallel Q_{X_1, X_2}) \underset{\substack{\uparrow \\ \text{item (iv) on DPI}}}{=} D_f(P_{X_1} \parallel Q_{X_1})$$

All in all : If $P_{x_2|x_1} = Q_{x_2|x_1}$ then

$$D_f(P_{x_1} \parallel Q_{x_1}) = D_f(P_{x_1, x_2} \parallel Q_{x_1, x_2}) = D_f(P_{x_1, x_2, Y} \parallel Q_{x_1, x_2, Y})$$

where

$$P_{x_1, x_2, Y} = P_{x_1} P_{x_2|x_1} P_{Y|x_1, x_2}$$

$$Q_{x_1, x_2, Y} = Q_{x_1} P_{x_2|x_1} P_{Y|x_1, x_2}$$

Examples of Transition Kernels

$$P_{Y|X}(\cdot | \cdot) : (X, \mathcal{F}) \rightarrow (Y, \mathcal{G})$$

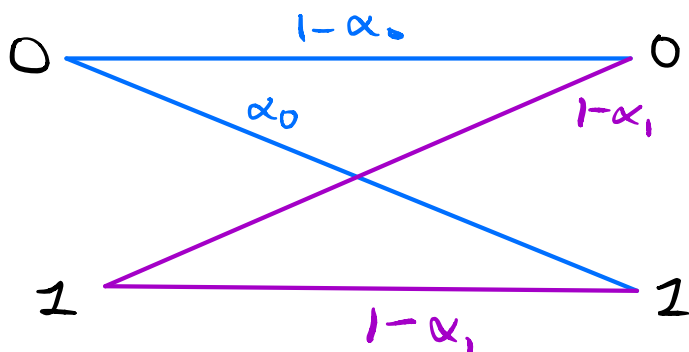
$$(i) P_{Y|X}(\cdot | x) \in \mathcal{P}(Y) \quad \forall x \in X$$

$$(ii) P_{Y|X}(B | \cdot) \text{ r.v. w.r.t } (X, \mathcal{F}) \quad \forall B \in \mathcal{G}$$

• Binary Channel

$$X = \{0, 1\}$$

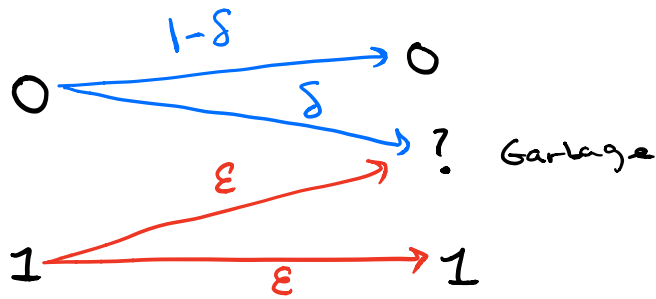
$$Y = \{0, 1\}$$



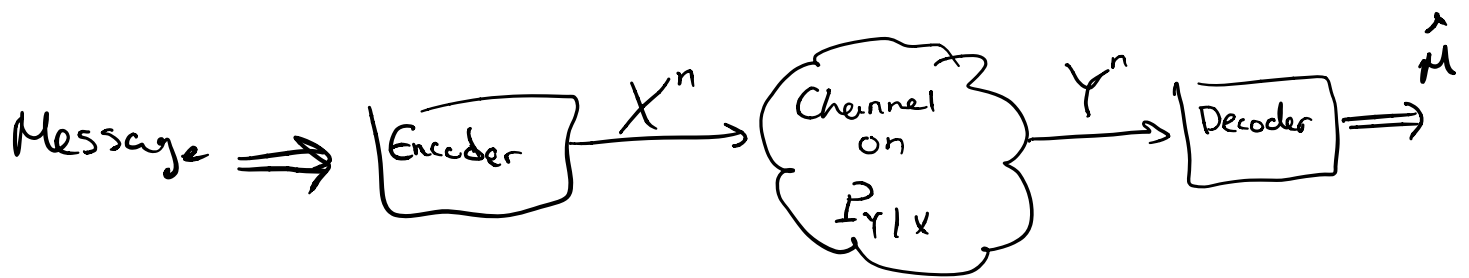
$$P_{Y|X}(\cdot | 0) = \text{Ber}(\alpha_0)$$

$$P_{Y|X}(\cdot | 1) = \text{Ber}(1 - \alpha_1)$$

- Binary Erasure Channel



Shannon's Channel Capacity Result



code pair of encoding + decoding function

$$C_n = (f_n, g_n)$$

Fix $C_n = (f_n, g_n)$

$$\Rightarrow P_{MX^n Y^n \hat{M}}^{(C_n)}$$