FANO'S INEQUALITY: A TWO-STEP PROOF

THEOREM: Let X, Y, Z be discrete random variables.

Define $A(z) = \sum_x \sum_y p(y) p(z|x, y)$.

Then: $H(X|Y) \le H(Z) + E[\log A]$.

(proof shown in class).

Corollary (Fano's Inequality): Let X, Y be random variables in the set $\{X_1, \ldots, X_r\}$. Let $P_e = P\{X \ne Y\}$. Then:

$$H(X|Y) \le H(P_e) + P_e \log(|\chi| - 1)$$

Proof: Define $Z = \left\{ \begin{array}{l} 0,\ X = Y \\ 1,\ X \ne Y \end{array} \right\}$. Note that $A(0) = A(z = 0) = 1$ and $A(1) = |\chi| - 1$ (see below). Then,

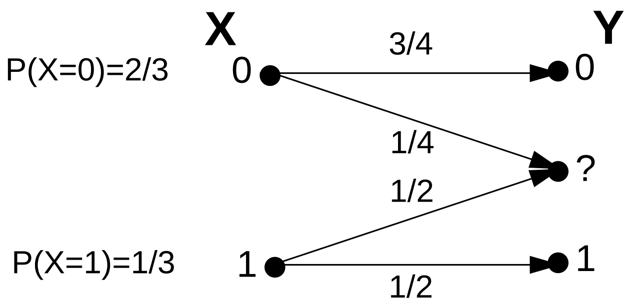
$$\begin{aligned} H(X|Y) &\le H(Z) + E[\log A] \\ &= H(P_e) + (1 - P_e)\log 1 + P_e \log(|\chi| - 1). \\ &= H(P_e) + P_e \log(|\chi| - 1) \end{aligned}$$

Note:

$$\begin{aligned} A(0) &= \sum_x \sum_y p(y) p(z = 0|x, y) \\ &= \sum_x \sum_y p(y) p(x = y|x, y) \\ &= \sum_y p(y) \sum_x p(x = y|x, y) \\ &= \sum_y p(y)(1) \\ &= 1 \end{aligned}$$

$$\begin{aligned} A(1) &= \sum_x \sum_y p(y) p(z = 1|x, y) \\ &= \sum_x \sum_y p(y) p(x \ne y|x, y) \\ &= \sum_y p(y) \sum_x p(x \ne y|x, y) \\ &= \sum_y p(y)(|\chi| - 1) \\ &= |\chi| - 1 \end{aligned}$$

Example: Discrete Memoryless Channel (Binary Erasure Channel)



Let’s start with some basics:

1. The “transition probabilities are given as input data:

$$P(Y=0|X=0) = 3/4 \qquad P(Y=0|X=1) = 0$$
$$P(Y=?|X=0) = 1/4 \qquad P(Y=?|X=1) = 1/2$$
$$P(Y=1|X=0) = 0 \qquad P(Y=1|X=1) = 1/2$$

We can write this compactly as:

$$\{p_{ij}\} = P(Y=Y_i|X=X_j) = \begin{bmatrix} 3/4 & 0 \\ 1/4 & 1/2 \\ 0 & 1/2 \end{bmatrix}$$

Note that the columns of this matrix sum to 1 (why?), but that the rows do not (why not?).

2. What about $p(y)$? (can be found several ways)

Let’s consider the case $P(Y=?)$ in detail:

$$P(Y=?) = P(X=0)P(Y=?|X=0) + P(X=1)P(Y=?|X=1) = \left(\frac{2}{3}\right)\left(\frac{1}{4}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{2}\right) = \frac{1}{3}.$$

Similarly,

$$P(Y=0) = P(X=0)P(Y=0|X=0) + P(X=1)P(Y=0|X=1) = \left(\frac{2}{3}\right)\left(\frac{3}{4}\right) = \frac{1}{2}$$

$$P(Y=1) = P(X=0)P(Y=1|X=0) + P(X=1)P(Y=1|X=1) = \left(\frac{1}{3}\right)\left(\frac{1}{2}\right) = \frac{1}{6}.$$

Hence,

$$\{P(Y=Y_i)\} = \begin{bmatrix} 1/2 & 1/3 & 1/6 \end{bmatrix}$$

Note that $\sum_y p(y) = 1$.

3. What about the joint distribution, $P(X = X_j, Y = Y_i)$. Consider $P(X = 0, Y = ?)$:

$$P(X = 0, Y = ?) = P\{X = 0\}P\{Y = ?|X = 0\} = (2/3)(1/4) = 1/6$$

   Hence, noting that rows correspond to X and colums to Y,

$$\{p(x, y)\} = \begin{bmatrix} 1/2 & 0 \\ 1/6 & 1/6 \\ 0 & 1/6 \end{bmatrix}$$

   Also, note that $\sum_x \sum_y p(x, y) = 1$, and that $p(y) = \sum_x p(x, y)$, which can be computed by summing the rows of $\{p(x, y)\}$. Similarly, $p(x) = \sum_y p(x, y)$, which can be computed by summing the columns of $\{p(x, y)\}$.

4. Next, we would like to compute $P(X = X_j|Y = Y_i)$:

$$P(X = 0|Y = ?) = P(X = 0,Y = ?)/P(Y = ?) = (1/6)/1/3 = 1/2$$

   Similarly,

$$\{P(X = X_j|Y = Y_i)\} = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \\ 0 & 1 \end{bmatrix}$$

   Note that the rows sum to 1, but the columns don't (why?).

Now we can begin our entropy-related calculations:

5. $H(X) = \sum_x p(x)\log\left(\frac{1}{p(x)}\right) = (2/3)\log\left(\frac{1}{(2/3)}\right) + (1/3)\log\left(\frac{1}{(1/3)}\right) = 0.918 \text{ bits}$

6. $H(Y) = \sum_y p(y)\log\left(\frac{1}{p(y)}\right) = 1.46 \text{ bits}$ Why is $H(Y) > H(X)$?

7. Next, let's calculate $H(Y|X)$:

$$H(Y|X) = \sum_x \sum_y p(x, y)\log(1/p(y|x)) = 0.87 \text{ bits}$$

8. What about the joint entropy:

$$H(X, Y) = \sum_x \sum_y p(x, y)\log(1/(p(x, y))) = 1.79 \text{ bits}$$

   But we could also calculate this using the chain rule:

$$H(X, Y) = H(X) + H(Y|X) = 0.918 + 0.87 = 1.79 \text{ bits}$$

9. Now, we want to compute the less intuitive $H(X|Y)$:

Method 1: Straightforward application of the basic equation

$$\begin{aligned} H(X|Y) &= \sum_x \sum_y p(x, y)\log(1/p(x|y)) \\ &= (1/2)\log(1) + (0)\log(0) + (1/6)\log(1/(1/2)) + (1/6)\log(1/(1/2)) + \\ &\quad + (0)\log(0) + (1/6)\log(1) \\ &= 1/3 \text{ bits} \end{aligned}$$

Method 2: Using intuition :)

$$\begin{aligned} H(X|Y = 0) &= 0 \\ H(X|Y = 1) &= 0 \\ H(X|Y = ?) &= \sum_x p(x|y)\log(1/(p(x|y))) \\ &= (1/2)\log(1/(1/2)) + (1/2)\log(1/(1/2)) \\ &= 1 \end{aligned}$$

and, $H(X|Y)$ is just the average of these:

$$\begin{aligned} H(X|Y) &= \sum_y p(y)H(X|Y = Y_i) \\ &= 1/3 \end{aligned}$$

10. Finally, from Fano's Inequality:

$$H(X|Y) \le H(P_e) + P_e\log(|\chi| - 1)$$

and,

$$H(X|Y) = \frac{1}{3}$$

An error occurs when $X \ne Y$, which only happens when $Y = ?$.
Recall that $|\chi| - 1 = 2$, $P_e = P(Y = ?) = 1/3$, and $1 - P_e = 2/3$, so:

$$H(P_e) = (1/3)\log(1/(1/3)) + (2/3)\log(1/(2/3)) = 0.918$$

and,

$$(1/3) \le 0.918 + (1/3)\log 2 = 1.25 \text{ bits}.$$

11. While we are at it, let's check a few more relationships involving mutual information:

$$I(X;Y) = H(X) - H(X|Y) = 0.92 - 0.33 = 0.59 \text{ bits}$$

$$I(X;Y) = H(Y) - H(Y|X) = 1.46 - 0.87 = 0.59 \text{ bits}$$

$$I(X;Y) = H(X) + H(Y) - H(X, Y) = 0.92 + 1.46 - 1.79 = 0.59 \text{ bits}$$