Exercise: Fun with entropies

(Source: Mackay.)

Consider the joint distribution $p(X, Y)$

		x			
		1	2	3	4
y	1	$1/8$	$1/16$	$1/32$	$1/32$
	2	$1/16$	$1/8$	$1/32$	$1/32$
	3	$1/16$	$1/16$	$1/16$	$1/16$
	4	$1/4$	0	0	0

1. What is the joint entropy $H(X, Y)$?
2. What are the marginal entropies $H(X)$ and $H(Y)$?
3. The entropy of X conditioned on a specific value of y is defined as

$$H(X|Y = y) = -\sum_x p(x|y) \log p(x|y) \tag{1}$$

   Compute $H(X|y)$ for each value of y. Does the posterior entropy on X ever increase given an observation of Y?

4. The conditional entropy is defined as

$$H(X|Y) = \sum_y p(y) H(X|Y = y) \tag{2}$$

   Compute this. Does the posterior entropy on X increase or decrease when averaged over the possible values of Y?

5. What is the mutual information between X and Y?