## Problem 1.

**a.** The entropy gets maximum when this random variable subjects to a uniform distribution in the interval [a, b]. It is easy to calculate it H = log(b-a).

**b.** The tighest lower bound is negative infinity, when the distribution becomes Dirac-delta function.

**c.** A tighter lower bound is –log(d), corresponding to a uniform distribution whose support is within [a, b] but the length of interval is 1/d.

**d.** The tightest upper bound will be infinity, which corresponds to a uniform distribution over an infinite support. The tightest lower bound is the same as in (c).

## Problem 2.

**a.**

```
function H=discreteEntropy(probs);

sum_prob=sum(probs);
    if sum_prob~=1
        error('Not a legal distribution');
    end;

H=0;
for n=1:length(probs)
    p=probs(n);
    if p~=0
        H = H - probs(n)*log2(probs(n));
    end
end
```

Example: A distribution with probabilities [.5 .25 .125 .125] for four events.
Output: 1.75

**b.**

```
function I=mutualInformation(prob1, prob2, joint_prob);

I=discreteEntropy(prob1)+discreteEntropy(prob2)-discreteEntropy(joint_prob(:));
```

**c.**

```
function D=KL(prob1,prob2)

D=0;
for i=1:length(prob1)
    if prob2(i)==0 & prob1(i)~=0
        disp('D is infinity');
        return;
    end
    if prob1(i)~=0 & prob2(i)~=0
        D=D+prob1(i)*(log(prob2(i))-log(prob1(i)));
    end
```

*end*

**Problem 3**

**a.** X and Y are not independent.

We know that: X and Y are independent iff P(X=x, Y=y) = P(X=x)*P(Y=y) for any x, and any y.

Without loss of generality, let X change over columns and and Y changes over rows.

Since that is a probability distribution in the figure, P(X=x) should be equal to the sum of all values in column x, and P(Y=y) the sum of all values in row y. If X and Y are independent, the value in the cell (x, y) should be equal to P(X=x)*P(Y=y). Check the figure, it is easy to see that it doesn't hold for many cells. For example, if one cell is black, which means P(X, Y) at the cell has value 0, at least one of P(X) and P(Y) should be 0, i.e. at least the whole row or the whole column the cell located in should black.

**4.**

**a.** The entropy of the distribution of Pixel values: 2.125. Then the number of bits of information in this image is 2.125x6=12.75.

**b.** Entropy goes up, because the distribution now is more close to a uniform distribution.

**c.** No, he didn't increase the amount of information in the image, because he randomly changed the values and also introduced more bins.

**d.** I think he does preserve more information. But I am not very clear about the reason…