Q 1.10.

$$H(p) = -\sum_{i=1}^{K} p_i \log p_i$$

for categorical distribution, the constraint should be:

$$\sum_{i=1}^{K} p_i = 1$$

To find the highest entropy, we use Lagrange multiplier

$$L(p, \lambda) = -\sum_{i=1}^{K} p_i \log(p_i) + \lambda\left(\sum_{i=1}^{K} p_i - 1\right)$$

$$\frac{\partial L}{\partial p_i} = \frac{\partial(-p_i \log p_i)}{\partial p_i} + \lambda$$

$$= -\log p_i - p_i \cdot \frac{1}{p_i} + \lambda$$

$$= -\log p_i - 1 + \lambda$$

Let $-\log p_i - 1 + \lambda = 0$, we get

$$\log p_i = \lambda - 1$$

This suggests for k categorical value, the probability for each category is the same, $p_i = \frac{1}{K}$

To confirm $p_i = \frac{1}{K}$ is the maximum entropy,

we compare $H(p_i = \frac{1}{K}) = \log K \geq 0$ with

$H(p_1 = 1, p_{i \neq 1} = 0) = 0$

So that when $p_i = \frac{1}{K}$, the categorical distribution has the highest entropy $\log K$