Exercise 12.13.15 For any $p \in (0,1)$ consider the *binary entropy function*

$$H(p) = -p \ln p - (1-p)\ln(1-p).$$

(a) Show that $H(p)$ is the entropy associated with a Bernoulli random variable.
(b) Verify the following relation between the derivative of the binary entropy and the logit function:

$$\frac{dH(p)}{dp} = -\ln\Big(\frac{p}{1-p}\Big).$$

(a) $x \in \{0, 1\}$: 이진 r.v

$p_X(x) = \begin{cases} p & \text{if } x=1 \\ 1-p & \text{if } x=0 \end{cases}$ 이므로 $p(x) = p^x \cdot (1-p)^{1-x}$

$H(x) = -\sum_x p(x|p) \cdot \ln p(x|p)$

$= -p(x=0|p) \cdot \ln p(x=0|p) - p(x=1|p) \cdot \ln(x=1|p)$

$= -(1-p) \cdot \ln(1-p) - p \cdot \ln p$

(b) $H(p) = -\ln(1-p) + p \cdot \ln(1-p) - p \cdot \ln p$

$\frac{dH(p)}{dp} = -(-\frac{1}{1-p}) + \ln(1-p) + p \cdot (-\frac{1}{1-p}) - \ln p - p \cdot \frac{1}{p}$

$= \frac{1}{1-p} - \frac{p}{1-p} - 1 + \ln\frac{1-p}{p}$

$= \ln(\frac{1-p}{p})$

$= -\ln(\frac{p}{1-p})$

Exercise 12.13.9 How does the capacity of a network change when:
(a) An extra fully-connected layer is added to the network;
(b) Some neurons are dropped out of the network;
(c) The weights are constrained to be kept small.

(a) capacity가 증가하나 session 내용에서 MNIST data에 대한 FC layer 추가에 따른 capacity의 증가의 limit이 있다고 하였으니 한계는 있을 것이라고 생각한다

(b) Drop out의 효과를 생각해 보았을 때에도_ 성능의 향상이 이루어질 것

(c) weight의 크기를 제한하여 regularization이 이루어지는 효과를 얻을 수 있으니 성능이 좋아질 것이다.