(1) $\|Xw - y\|^2$

$$= (Xw - y)^T (Xw - y)$$
$$= (w^T X^T - y^T)(Xw - y)$$
$$= w^T X^T X w - w^T X^T y - y^T X w - y^T y$$

$$\nabla_w \|Xw - y\|^2$$

$$= \nabla_w (w^T X^T X w - w^T X^T y - y^T X w - y^T y)$$
$$= 2X^T X w - 2X^T y$$

$$2X^T X w - 2X^T y = 0$$
$$\Rightarrow 2X^T (Xw - y) = 0$$
$$\Rightarrow X^T X w - X^T y = 0$$
$$\Rightarrow w = (X^T X)^{-1} X^T y$$

(2) $\|Xw - y\|^2 + \lambda \|w\|^2$

$$= (Xw - y)^T (Xw - y) + \lambda w^T w$$
$$= (w^T X^T - y^T)(Xw - y) + \lambda w^T w$$
$$= w^T X^T X w - w^T X^T y - y^T X w + y^T y + \lambda w^T w$$

$$\nabla_w (\|Xw - y\|^2 + \lambda \|w\|^2)$$

$$= \nabla_w \left(w^T x^T x w - w^T x^T y - y^T x w + y^T y + \lambda w^T w \right)$$

$$= 2 x^T x w - 2 x^T y + 2 \lambda w$$

$$\nabla_w = 0$$

$$\Rightarrow 2 x^T x w - 2 x^T y + 2 \lambda w = 0$$

$$\Rightarrow w \left(x^T x + \cancel{2} \lambda I \right) = x^T y$$

$$\Rightarrow w = x^T y \left(x^T x + \lambda I \right)^{-1}$$

2 probability $= P(H)P(H)P(T)P(T)P(H)$

$= p^2 (1-p)^2 p$

$= p^3 (1+p^2-2p)$

$= p^3 + p^5 - 2p^4$

$\log_p$ (Probability)

$= \log_e (p^3 + p^5 - 2p^4)$

$= \cancel{3\log_e (1-p)^2}\ \log_e p^3 (1-p)^2$

$= 3\log_e p + 2\log_e (1-p)$

2 2

(a) Probability of choosing 1st coin

$= 1/2$, $P(T) = 1 - \{P(H) = 1/2$

Probability $= \frac{1}{2} P(H) P(H) P(T) P(T) P(H)$

$= \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}$

$= \frac{1}{2^6}$

(b) $P(H) = 2/3$, $P(T) = 1/3$

Probability of choosing 2nd coin $= \frac{1}{2}$

Probability $= \frac{1}{2} P(H) P(H) P(T) P(T) P(H)$

$= \frac{1}{\cancel{2}} \times \frac{\cancel{2}}{3} \times \frac{2}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3}$

$= 4/243$

(3) $pr(H) = p$

$Pr(T) = 1-p$

Probability = $P(H)\,P(H)\,P(T)\,P(T)\,P($

$= p^3 (1-p)^2$

$\log(\text{Probability}) = \log_e p^3 (1-p)^2$

$= \log_e p^3 + \log_e (1-p)^2$

$= 3\log_e p + 2\log_e (1-p)$

maximize the $\log_e(\text{probability})$

$\nabla_p (\log_e(\text{Probability})) = 0$

$\Rightarrow \nabla_p (3\log_e p + 2\log_e(1-p)) = 0$

$\Rightarrow \frac{d}{dp} 3\log_e p + \frac{d}{dp} 2\log_e(1-p) = 0$

$\Rightarrow 3/p - \frac{2}{1-p}(-1) = 0$

$\Rightarrow 3/p + 2/1-p = 0$

$\Rightarrow 3 - 3p = 2p \Rightarrow 3 = 5p$

$\Rightarrow p = 3/5$