Exercise 1

With 2 classes C_1 and C_2, the posterior probability for C_1 can be written as:

$$p(C_1|x) = \frac{P(x|C_1)p(C_1)}{P(x|C_1)p(C_1) + P(x|C_2)p(C_2)}$$
$$= \frac{1}{1 + \frac{P(x|C_2)p(C_2)}{P(x|C_1)p(C_1)}}$$
$$= \frac{1}{1+e^{-a}}\ with\ a = log\left(\frac{P(x|C_1)p(C_1)}{P(x|C_2)p(C_2)}\right)$$
$$= \sigma(a)$$

With $\hat{y} = w_0 + w_1\phi(x_1) + w_2\phi(x_2) + \ldots + w_n\phi(x_n) = w^T\phi$, the model of logistic regression is defined as

$$p(C_1|\phi) = y(\phi) = \sigma\left(w^T\phi\right)$$
$$p(C_2|\phi) = 1 - p(C_1|\phi)$$

The likelihood function is

$$p(t|w) = \prod y_n^{t_n}(1-y_n)^{1-t_n}$$

where $t = (t_1, \ldots t_n)^T\ and\ y_n = p(C_1|\phi_n)$

We defined the cross entropy function as

$$L = -log\ p(t|w)$$
$$= -\sum t_n log(y_n) - \sum (1-t_n)log(1-y_n)$$
$$= -t^T log(y) - (1-t)^T log(1-y)$$

Taking the gradient of the function with respesct to w

$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial y} \times \frac{\partial y}{\partial z} \times \frac{\partial z}{\partial w}\ where\ z = w^T\phi$$
$$\frac{\partial L}{\partial y} = \left[\frac{-t^T}{y} + \frac{(1-t)^T}{1-y}\right] = \frac{y-t}{y(1-y)}$$
$$\frac{\partial y}{\partial z} = \frac{\partial \sigma(z)}{\partial z} = \left(\frac{1}{1+e^{-z}}\right)' = \sigma(z)(1-\sigma(z)) = y(1-y)$$
$$\frac{\partial z}{\partial w} = \phi$$
$$\leftrightarrow \frac{\partial L}{\partial w} = \frac{y-t}{y(1-y)} \times y(1-y) \times \phi = (y-t)\phi$$

Exercise 2

$$f'(x) = f(x)[1 - f(x)] \; with \; 0 < f(x) < 1$$
$$\frac{f'(x)}{f(x)[1 - f(x)]} = 1$$
$$\int \frac{f'(x)}{f(x)[1 - f(x)]} dx = \int 1 dx$$
$$\int \frac{d(f(x))}{f(x)[1 - f(x)]} = x + C$$
$$\int \frac{1}{f(x)} + \frac{1}{1 - f(x)} d(f(x)) = x + C$$
$$ln|f(x)| - ln|1 - f(x)| = x + C$$
$$\frac{f(x)}{1 - f(x)} = e^{x+C}$$
$$f(x) = e^{x+C} - f(x)e^{x+C}$$
$$f(x) = \frac{e^{x+C}}{1 + e^{x+C}} = \frac{e^x}{C + e^x}$$