hw1.

1. $y(x, w) = w_0 + w_1 x + w_2 x^2 + \cdots + w_M x^M = \sum_{j=0}^{M} w_j x^j$ ①

$E(w) = \frac{1}{2} \sum_{n=1}^{N} \{y(x_n, w) - t_n\}^2$ ②

Substituting ① into ②. and differentiating with respect to w_i

$\sum_{n=1}^{N} \left(\sum_{j=0}^{M} w_j x_n^j - t_n \right) x_n^i = 0.$ ③.

by re-arrange function ③. we get $\{w_j\}$ that minimize the sum-of-squares error function

2. $p(apple) = p(apple|red)\,p(red) + p(apple|blue)\,p(blue) + p(apple|green)\,p(green)$

$= \frac{3}{10} \times 0.2 + \frac{1}{2} \times 0.2 + \frac{3}{10} \times 0.6$

$= 0.34$

$p(green|orange) = \frac{p(orang|green)\,p(green)}{p(orang)}$

$p(orange) = p(orang|red)\,p(red) + p(orang|blue)\,p(blue) + p(orange|green)\,p(green)$

$= \frac{4}{10} \times 0.2 + \frac{1}{2} \times 0.2 + \frac{3}{10} \times 0.6$

$= 0.36$

$p(green|orange) = \frac{3}{10} \times \frac{0.6}{0.36} = 0.5$

x and z are independent, $p(x, z) = p(x)\,p(z)$, so.

$E[x+z] = \iint (x+z)\,p(x)\,p(z)\,dx\,dz$

$= \int x\,p(x)\,dx + \int z\,p(z)\,dz.$

$= E[x] + E[z].$

$(x+z - E[x+z])^2 = (x - E[x])^2 + (z - E[z])^2 + \underbrace{2(x - E[x])(z - E[z])}_{\text{integrate to zero with } p(x)p(z)}$

$var[x+z] = \iint (x+z - E[x+z])^2\, p(x)\,p(z)\,dx\,dz.$

$= \int (x - E[x])^2 p(x)\,dx + \int (z - E[z])^2 p(z)\,dz.$

$= var[x] + var[z]$

4.1

① $L(x_1, x_2, x_3, \cdots x_n|\lambda) = \prod^{n} (\frac{\lambda^{x_1} e^{-\lambda}}{x_1!}) + \prod^{n} (\frac{\lambda^{x_2} e^{-\lambda}}{x_2!}) + \cdots + \prod^{n} (\frac{\lambda^{x_n} e^{-\lambda}}{x_n!})$.

$$\cancel{L(\lambda) = -\lambda(\ln \frac{\lambda^{x_1}}{n x_1!} + \ln \frac{\lambda^{x_2}}{n x_2!} + \cdots + \ln \frac{\lambda^{x_n}}{n x_n!})}$$

$$\cancel{= \ln \frac{\lambda^{x_1+1}}{n x_1!}}$$

$$\cancel{= \ln(\lambda^{x_1} - \lambda) - \ln(n x_1!) + \ln(\lambda^{x_2} - \lambda) - \ln(n x_2!) + \cdots + \ln(\lambda^{x_n} - \lambda) -}$$

$$= e^{-n\lambda} \cdot \prod^{n} (\frac{\lambda^{x_i}}{x_i!})$$

$$\ln L = -n\lambda + \sum_{i=1}^{n} (x_i \ln \lambda - \ln x_i).$$

$$\frac{d\ln L}{d\lambda} = -n + \sum_{i=1}^{n} \frac{x_i}{\lambda} = 0.$$

$$\lambda = \frac{1}{n} \sum_{i=1}^{n} x_i$$

4.2

$$L(x_i; \lambda) = \prod_{i=1}^{n} \frac{1}{\lambda} e^{-\frac{x}{\lambda}}$$

$$\ln(L) = e^{-\frac{\sum_{i=1}^{n} x_i}{\lambda}} \cdot \prod_{i=1}^{n} \frac{1}{\lambda}.$$

$$= -\frac{\sum_{i=1}^{n} x_i}{\lambda} + \ln(\frac{1}{\lambda^n}) = -\frac{\sum_{i=1}^{n} x_i}{\lambda} + n\ln(\frac{1}{\lambda})$$

$$\frac{\ln(L)}{d\lambda} = \lambda^{-2} \sum_{i=1}^{n} x_i + n\lambda \cdot (-\frac{1}{\lambda^2}) = 0$$

$$\lambda^{-2} \sum_{i=1}^{n} x_i - \frac{n}{\lambda} = 0$$

$$\sum_{i=1}^{n} x_i - n\lambda = 0.$$

$$\sum_{i=1}^{n} x_i = n\lambda.$$

$$\lambda = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$S_{(a)}$ $p(\text{mistake}) = p(x \in R_1, C_2) + p(x \in R_2, C_1)$

$= \int_{R_1} p(x, C_2)\,dx + \int_{R_2} p(x, C_1)\,dx$

$p(\text{correct}) = \sum_{k=1}^{K} p(x \in R_k, C_k)$

$= \sum_{k=1}^{K} \int_{R_k} p(x, C_k)\,dx$

here $p(\text{correct}) = \int_{R_1} p(x, C_1)\,dx + \int_{R_2} p(x, C_2)\,dx$

b) $y(x) = E_t[t|x]$.

$$E[L(t, y(x))] = \iint \|y(x) - t\|^2 p(x,t)\,dx\,dt.$$

$$= \iint \|y(x) - E_t[t|x] + E_t[t|x] - t\|^2 p(x,t)\,dx\,dt.$$

$$\cancel{= \iint \|E_t[t|x] - t\|\, p(x,t)\,dx\,dt.}$$

~~this is to minimize~~

$$\|y(x) - t\|^2 = \|y(x) - E_t[t|x]\|^2 + 2\|y(x) - E_t[t|x]\| \cdot \|E_t[t|x] - t\| + \|E_t[t|x] - t\|^2$$

the cross-term vanishes so we need to keep

$\|y(x) - E_t[t|x]\|^2$ to be minimized.

this be minimized when $y(x) = E_t[t|x]$

6. a) $H[X] = -\sum_i p(x_i) \ln p(x_i)$ X is continuous. so $H[X] = -\int p(x) \ln p(x)\,dx$

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad H(X) = -\int \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \ln\left(\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right) dx$$

$$= \frac{1}{2}\{1 + \ln(2\pi\sigma^2)\}.$$

b) $I[x,y] \equiv KL(p(x,y) \| p(x)p(y)) = -\int p(x,y) \ln p(x)p(y)\,dx - \left(-\int p(x,y) \ln p(x,y)\,dx\right)$

$$= -\iint p(x,y) \ln\left(\frac{p(x)p(y)}{p(x,y)}\right) dx\,dy \qquad I(x,y) \geq 0.$$

$H[y|x] = -\iint p(y,x) \ln p(y|x)\,dy\,dx \qquad H[x,y] = H[y|x] + H[x]$

$I[x,y] = H[x] - H[x|y] = H[y] - H[y|x]$.

$$I[y|x] = KL(p(y,x) \| p(y)p(x)) = -\iint p(y,x) \ln \frac{p(y)p(x)}{p(y,x)}\,dy\,dx.$$

$$= H[y] - H[y|x] = H[x] - H[x|y].$$