1201[illegible] [illegible]

Q1: Answer: $y(x, \vec{w}) = w_0 + w_1 x + w_2 x + \cdots + w_M x^M = \sum_{j=0}^{M} w_j x^j$

$$E(\vec{w}) = \frac{1}{2} \sum_{n=0}^{N} \{y(x_n, \vec{w}) - t_n\}^2 = \frac{1}{2} \sum_{n=0}^{N} \{y^2(x_n, \vec{w}) + t_n^2 - 2y(x_n, \vec{w}) t_n\}$$

$$\frac{\partial E(\vec{w})}{\partial \vec{w}} = \frac{1}{2} \sum_{n=0}^{N} \left\{2y(x_n, \vec{w}) \frac{\partial y(x_n, \vec{w})}{\partial \vec{w}} - 2t_n \frac{\partial y(x_n, \vec{w})}{\partial \vec{w}}\right\} = \sum_{n=0}^{N} \left\{(y(x_n, \vec{w}) - t_n) \frac{\partial y(x_n, \vec{w})}{\partial \vec{w}}\right\}$$

when the sum-of-squares error function is minimal, $\frac{\partial E(\vec{w})}{\partial \vec{w}} = 0$

$$\Rightarrow \sum_{n=0}^{N} \left\{[y(x_n, \vec{w}) - t_n] \frac{\partial y(x_n, \vec{w})}{\partial \vec{w}}\right\} = 0 \Rightarrow \sum_{n=0}^{N} \left\{\left(\sum_{j=0}^{M} x_n^j\right) \cdot \left(\sum_{j=0}^{M} w_j x_n^j - t_n\right)\right\} = 0.$$

let $\vec{w} = \{w_j\}$, $\vec{x} = \{x^j\}$ $\Rightarrow \sum_{n=0}^{N} \left\{\left(\sum_{j=0}^{M} x_n^j\right) (\vec{w} \cdot \vec{x}^T - t_n)\right\} = 0$

Q2:

	apples	oranges	limes
r	3	4	3
b	1	1	0
g	3	3	4

a): let p denotes the probability of selecting an apple.

$$p = p(r) \cdot p(\text{apples}|r) + p(b)\, p(\text{apples}|b) + p(g)\, p(\text{apples}|g)$$
$$= 0.2 \times \frac{3}{10} + 0.2 \times \frac{1}{2} + 0.6 \times \frac{3}{10} = 0.34.$$

(b): orange, green) = p(orange| 1) · p(green) = p(green|orange) · p(orange)

$$p(\text{orange}, \text{green}) = p(\text{orange}|\text{green})\, p(\text{green}) = \frac{3}{10} \times 0.6 = 0.18.$$

$$p(\text{orange}) = p(r)\, p(\text{orange}|r) + p(b)\, p(\text{orange}|b) + p(g)\, p(\text{orange}|g)$$
$$= 0.2 \times \frac{4}{10} + 0.2 \times \frac{1}{2} + 0.6 \times \frac{3}{10} = 0.36.$$

$$\Rightarrow p(\text{green}|\text{orange}) = \frac{p(\text{orange}, \text{green})}{p(\text{orange})} = \frac{0.18}{0.36} = 0.5$$

Q3: $E[x+z] = \cancel{\int p(x+z)(x+z)\, d(x+z)} = \int_x \int_z p(X=x, Z=z)(x+z)\, d(x+z)$

$\because$ x and z are independent variables $\therefore$ $p(X=x, Z=z) = p(X=x) \cdot p(Z=z)$

$$\Rightarrow E[x+z] = \int_x \int_z p(X=x, Z=z)(x+z)\, d(x+z) = \int_x \int_z p(X=x)\, p(Z=z)(x+z)\, dx\, dz$$

$$= \int_x \int_z p(X=x) p(Z=z) x\, dx\, dz + \int_x \int_z p(X=x)\, p(Z=z) z\, dx\, dz = \int_x p(X=x) \cdot x \int_z p(Z=z)\, dz\, dx$$

$$+ \int_z p(Z=z) \cdot z \cdot dz \int_x p(x)\, dx = \int_x p(X=x) \cdot x\, dx \cdot 1 + \int_z p(Z=z) \cdot z\, dz \cdot 1 = E(x) + E(z).$$

$\mathrm{Var}[X+Z] = E[((X+Z) - E[X+Z])^2] = E[((X+Z) - (E[X]+E[Z]))^2]$

$= E[(X+Z)^2 + (E[X]+E[Z])^2 - 2(X+Z)(E[X]+E[Z])]$ $\cancel{= E[X^2+Z^2+2XZ+E^2[X]+E^2[Z]+2E[X]E[Z]}$

$= E(X^2+Z^2+2XZ+E^2[X]+E^2[Z]+2E[X]E[Z] - 2XE[X] - 2ZE[X] - 2XE[Z] - 2ZE[Z])$

$= E[(X-E[X])^2 + (Z-E[Z])^2 + 2XZ + 2E[X]E[Z] - 2ZE[X] - 2XE[Z]]$

$= E[(X-E[X])^2] + E[(Z-E[Z])^2] + E[2XZ] + E(2E[X]E[Z]) - E[2ZE[X]] - E[2XE[Z]]$

$= \mathrm{Var}(X) + \mathrm{Var}(Z) + E[2XZ] + E[2E[X]E[Z]] - E[2ZE[X]] - E[2XE[Z]]$

$\because$ X and Z are independent variables

$\therefore$ $E[2XZ] = 2E[X]E[Z]$

$\therefore$ $\mathrm{Var}[X+Z] = \mathrm{Var}[X] + \mathrm{Var}[Z] + 2E[X]E[Z] + 2E[X]E[Z] - 2E[X]E[Z] - 2E[Z]E[X]$

$= \mathrm{Var}[X] + \mathrm{Var}[Z]$.

Q4: Answer. a): $p(x|\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$ $E(X) = \int_X p(x|\lambda)\, x \cdot dx = \lambda \cdot \int_x \frac{\lambda^x e^{-\lambda}}{x!} \cdot x \, dx$

$D = \{x_1, x_2, \cdots, x_n\}$, $p(\vec{x}|\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$,

$\ln p(\vec{x}|\lambda) = \sum_{i=1}^{n} \ln \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = \sum_{i=1}^{n} [\ln \lambda^{x_i} + \ln e^{-\lambda} - \ln(x_i!)] = \sum_{i=1}^{n} [x_i \ln\lambda - \lambda - \ln(x_i!)]$

$\frac{\partial[\ln p(\vec{x}|\lambda)]}{\partial \lambda} = \sum_{i=1}^{n} (\frac{x_i}{\lambda} - 1) = 0 \Rightarrow n\hat{\lambda} = \sum_{i=1}^{n} x_i \Rightarrow \hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} x_i$

b) $f(x) = \begin{cases} \frac{1}{\lambda} e^{-\frac{x}{\lambda}}, & x > 0 \\ 0, & x \le 0. \end{cases}$ $p(\vec{x}) = \prod_{i=1}^{n} \frac{1}{\lambda} e^{-\frac{x_i}{\lambda}}$

$\ln p(\vec{x}) = \sum_{i=1}^{n} [\ln(\frac{1}{\lambda}) + \ln e^{-\frac{x_i}{\lambda}}] = \sum_{i=1}^{n} (-\ln\lambda - \frac{x_i}{\lambda})$.

$\frac{\partial[\ln p(\vec{x}|\lambda)]}{\partial \lambda} = \sum_{i=1}^{n} (\frac{-1}{\lambda} + \frac{x_i}{\lambda^2}) = 0 \Rightarrow \hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

Q5: Answer: a). $p(\text{correct}) = p(\vec{x} \in R_1, C_1) + p(\vec{x} \in R_2, C_2) = \int_{R_1} p(\vec{x}, C_1)\, d\vec{x} + \int_{R_2} p(\vec{x}, C_2)\, d\vec{x}$.

$p(\text{mistake}) = p(\vec{x} \in R_1, C_2) + p(\vec{x} \in R_2, C_1) = \int_{R_1} p(\vec{x}, C_2)\, d\vec{x} + \int_{R_2} p(\vec{x}, C_1)\, d\vec{x}$.

b) $E[L(\vec{t}, \vec{y}(\vec{x}))] = \int\int ||\vec{y}(\vec{x}) - \vec{t}||^2 p(\vec{x}, \vec{t})\, d\vec{x}\, d\vec{t}$

let $L(\vec{x}, \vec{t}) = ||\vec{y}(\vec{x}) - \vec{t}||^2 = \cancel{\vec{y}(\vec{x})\vec{y}^T(\vec{x})} ||\vec{y}(\vec{x}) - E_{\vec{t}}[\vec{t}|\vec{x}] + E_{\vec{t}}[\vec{t}|\vec{x}] - \vec{t}||^2$

$= ||\vec{y}(\vec{x}) - E_{\vec{t}}[\vec{t}|\vec{x}]||^2 + 2(\vec{y}(\vec{x}) - E_{\vec{t}}[\vec{t}|\vec{x}])(E_{\vec{t}}[\vec{t}|\vec{x}] - \vec{t})^T + ||E_{\vec{t}}[\vec{t}|\vec{x}] - \vec{t}||^2$

$= L_1 + 2L_2 + 2L_3$

$\Rightarrow E[L(\vec{t}, \vec{y}(\vec{x}))] = \cancel{L_1 + 2L_2 + 2L_3}\ E_{\vec{x},\vec{t}}[L_1] + 2E_{\vec{x},\vec{t}}[L_2] + E_{\vec{x},\vec{t}}[L_3]$

$E_{\vec{x},\vec{t}}[L_1] = E_{\vec{x},\vec{t}}[||\vec{y}(\vec{x}) - E_{\vec{t}}[\vec{t}|\vec{x}]||^2] = E_{\vec{x}}[\ E_{\vec{t}}[(||\vec{y}(\vec{x}) - E_{\vec{t}}[\vec{t}|\vec{x}]||^2)|\vec{x}]\ ]$

$= E_{\vec{x}}[||\vec{y}(\vec{x}) - E_{\vec{t}}[\vec{t}|\vec{x}]||^2]$.

$E_{\vec{x},\vec{t}}[L_2] = E_{\vec{x},\vec{t}}[(\vec{y}(\vec{x}) - E_{\vec{t}}[\vec{t}|\vec{x}])(E_{\vec{t}}[\vec{t}|\vec{x}] - \vec{t})^T] = E_{\vec{x}}[E_{\vec{t}}[\{(\vec{y}(\vec{x}) - E_{\vec{t}}[\vec{t}|\vec{x}])(E_{\vec{t}}[\vec{t}|\vec{x}] - \vec{t})^T\}|\vec{x}]]$

$= E_{\vec{x}}[(\vec{y}(\vec{x}) - E_{\vec{t}}[\vec{t}|\vec{x}])\, E_{\vec{t}}[(E_{\vec{t}}[\vec{t}|\vec{x}] - \vec{t})^T|\vec{x}]]$.

$E_{\vec{t}}[(E_{\vec{t}}[\vec{t}|\vec{x}] - \vec{t})^T|\vec{x}] = E_{\vec{t}}[E_{\vec{t}}[\vec{t}|\vec{x}]|\vec{x}] - E_{\vec{t}}[\vec{t}|\vec{x}] = E_{\vec{t}}[\vec{t}|\vec{x}] - E_{\vec{t}}[\vec{t}|\vec{x}] = 0$

$\Rightarrow E_{\vec{x},\vec{t}}[L_2] = 0$.

$E_{\vec{x},\vec{t}}[L_3] = E_{\vec{x},\vec{t}}[||E_{\vec{t}}[\vec{t}|\vec{x}] - \vec{t}||^2] = E_{\vec{x}}[E_{\vec{t}}[||E_{\vec{t}}[\vec{t}|\vec{x}] - \vec{t}||^2|\vec{x}]] = E_{\vec{x}}[Var_{\vec{t}}[\vec{t}|\vec{x}]]$

$\Rightarrow E[L(\vec{t}, \vec{y}(\vec{x}))] = \int\int ||\vec{y}(\vec{x}) - E_{\vec{t}}[\vec{t}|\vec{x}]||^2 p(\vec{x})\, d\vec{x} + \int_{\Omega}\int Var_{\vec{t}}[\vec{t}|\vec{x}]\, p(\vec{x})\, d\vec{x}$.

minimize $E[L(\vec{t}, \vec{y}(\vec{x}))] \Rightarrow ||\vec{y}(\vec{x}) - E_{\vec{t}}[\vec{t}|\vec{x}]||^2 = 0 \Rightarrow \vec{y}(\vec{x}) = E_{\vec{t}}[\vec{t}|\vec{x}]$

Q6: (a). $H[x] = -\int p(x)\ln p(x)\, dx = -\int \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \ln\left[\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right] dx$

$= \int \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \ln(\sqrt{2\pi}\sigma)\, dx + \int \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \frac{(x-\mu)^2}{2\sigma^2}\, dx$

$= \ln(\sqrt{2\pi}\sigma)\int p(x)\, dx + \frac{1}{2\sigma^2}\int p(x)(x-\mu)^2\, dx = \cancel{\ln}\ \frac{1}{2}\ln(2\pi\sigma^2) + \frac{1}{2} = \frac{1}{2}\ln(2\pi e\sigma^2)$.

(b) $I(X;Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y)\log\frac{p(x,y)}{p(x)p(y)} = \sum_{x\in X}\sum_{y\in Y} p(x,y)(\log p(x,y)) - \sum_{x\in X}\sum_{y\in Y} p(x,y)\log p(x) - \sum_{x\in X}\sum_{y\in Y} p(x,y)\log p(y)$

$= \sum_{x\in X}\sum_{y\in Y} p(x,y)\log\frac{p(x,y)}{p(x)} - \sum_{y\in Y} p(y)\log p(y) = H(Y) + \sum_{x\in X}\sum_{y\in Y} p(x,y)\log p(y|x) = H(Y) - -\sum_{x\in X}\sum_{y\in Y} p(y|x)p(x)\log p(y|x)$

$= H(Y) - H(Y|X) = H(X) - H(X|Y)$