Question 2.

a) likelihood $= \prod_{i=1}^{n} \theta^{x_i} (1-\theta)^{(1-x_i)}$

log likelihood $= \log \prod_{i=1}^{n} \theta^{x_i} (1-\theta)^{(1-x_i)}$

$$= \sum_{i=1}^{n} x_i (\log\theta) + (1-x_i) \log(1-\theta)$$

$$= \sum_{i=1}^{n} x_i \log\theta + (n - \sum_{i=1}^{n} x_i) \log(1-\theta)$$

$$\frac{dll}{d\theta} = \sum_{i=1}^{n} x_i \frac{1}{\theta} + (n - \sum x_i)\frac{-1}{1-\theta} \overset{set}{=} 0$$

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

b) $\Pr(x; \theta_1 \cdots \theta_K) = \prod_{k=1}^{K} \theta_k^{x_k}$

likelihood : $L(\theta; \{x^{(i)}\}_{i=1}^{N}) = \prod_{i=1}^{n} \prod_{k=1}^{K} \theta_k^{x_{ki}}$

log likelihood $= \ln L(\theta; \{x^{(i)}\}_{i=1}^{N})$

$$\sum_{i=1}^{n} \log\left(\prod_{k=1}^{K} \theta_k^{x_{ki}}\right)$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{K} x_{ki} (\log \theta_k)$$

Lagrangian :

$$\ln L_\lambda(\theta; \{x^{(i)}\}_{i=1}^{N}) = \ln L(\theta; \{x^{(i)}\}_{i=1}^{N} + \lambda(1 - \sum_{k=1}^{K}\theta_k)$$

$$= \sum_{i=1}^{n} x_{ik} (\log\theta_k) + \lambda(1 - \sum_{k=1}^{K}\theta_k)$$

$$\frac{d \ln L_\lambda}{d\theta_k} = \sum_{i=1}^{n} x_{ik} \frac{1}{\theta_k} - \lambda. \overset{set}{=} 0$$

$$\hat{\theta_k} = \frac{1}{\lambda} \sum_{i=1}^{n} x_{ik}.$$

$$\frac{d \ln L_\lambda}{d\lambda} = 1 - \sum_{i=1}^{k} \theta_k \overset{set}{=} 0$$

$$\sum_{i=1}^{k} \theta_k = 1$$

So. $\sum_{i=1}^{k} \frac{1}{\lambda} \sum_{i=1}^{n} x_{ik} = 1$

$$\frac{1}{\lambda} \sum_{i=1}^{K} \sum_{i=1}^{n} x_{ik} = 1$$

$(* \sum_{k=1}^{K} \theta_k = 1)$ $\qquad \frac{1}{\lambda} \sum_{i=1}^{n} 1 = 1$

$$\hat{\lambda} = N$$

So $\hat{\theta} = \frac{1}{N} \sum_{i=1}^{n} x_{ik}.$

(c) $p(x; \mu, \sigma^2) = \frac{1}{(2\pi)^{\frac{d}{2}} \sqrt{\det(\Sigma)}} \exp\left(\frac{-(x-\mu)^T \Sigma^{-1} (x-\mu)}{2}\right)$

$$Lp(x; \mu, \Sigma) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{\frac{d}{2}} \sqrt{\det(\Sigma)}} \exp\left(\frac{-(x_i-\mu)^T \Sigma^{-1} (x_i-\mu)}{2}\right)$$

$\ln L\, p(x; \mu\, \Sigma)$

$$= -n\frac{d}{2} \ln(2\pi) - n\frac{1}{2} \ln(\det(\Sigma)) - \sum_{i=1}^{N} \frac{(x_i-\mu)^T \Sigma^{-1} (x_i-\mu)}{2}$$

$$\frac{d \ln L}{d\mu} = -\sum_{i=1}^{N} \frac{1}{2} \cdot 2(x_i - \mu) \cdot \Sigma^{-1}$$

$$= -\Sigma^{-1} \sum_{i=1}^{N} (x_i - \mu) \overset{set}{=} 0$$

$$\hat{u} = \frac{1}{N} \sum_{i=1}^{N} x_i = \bar{x}$$

* • the trace is invariant under cyclic permutation of matrix product

$$tr[ACB] = tr[CAB] = tr[BCA]$$

• Since $x^T A x$ is scalar, we can take its trace :

$$x^t A x = tr[x^T A x] = tr[x^T x A]$$

• $\frac{d}{dA} tr[AB] = B^T$

• $\frac{d}{dA} \log|A| = A^{-T}$

From above properties.

$$\frac{d}{dA} x^t A x = \frac{d}{dA} tr[x^T x A] = [x^T x]^T$$

$$= x^{TT} x = x x^T$$

$\frac{d}{d\Sigma^{-1}} \ln l(x; \mu, \Sigma)$

$$= \frac{n}{2} \Sigma - \frac{1}{2} \sum_{i=1}^{N} (x^i - \mu)(x^i - \mu)^T$$

$$= \frac{n}{2} \Sigma - \frac{1}{2} \sum_{i=1}^{N} (x_i - \mu)^T (x_i - \mu) \overset{set}{=} 0$$

$$n\Sigma = \sum_{i=1}^{N} (x_i - \mu)^T (x_i - \mu)$$

$$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^{N} (x_i - \mu)^T (x_i - \mu)$$

$$\frac{d \det(A)}{dA} = \det(A) A^{-T}$$

Question 3

a) $p(x; x_0, \gamma) = \dfrac{1}{\pi \exp(\gamma) [1 + (\frac{x-x_0}{\exp(\gamma)})^2]}$

$$\int_{-\infty}^{\infty} p(x; x_0, \gamma)\, dx$$

let $\dfrac{x-x_0}{\exp(\gamma)} = u$

$\dfrac{du}{dx} = \dfrac{1}{\exp(\gamma)}$ $\quad du = \dfrac{1}{\exp(\gamma)} dx$

$$= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{\exp(\gamma)(1+u^2)} dx$$

$$= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{1+u^2} du$$

$$= \frac{1}{\pi} \cdot \tan^{-1} \Big]_{-\infty}^{\infty}$$

$$= \frac{1}{\pi} \cdot \pi = 1$$

b) $E_{\exp(x; x_0, \gamma)}[x]$

$$= \int_{-\infty}^{\infty} x \cdot \frac{1}{\pi \exp(\gamma) [1 + (\frac{x-x_0}{\exp(\gamma)})^2]} dx$$

$$= \frac{1}{\pi \exp(\gamma)} \int_{-\infty}^{\infty} \frac{x-x_0}{1+(\frac{x-x_0}{\exp(\gamma)})^2} + \frac{x_0}{1+(\frac{x-x_0}{\exp(\gamma)})^2} dx$$

since $\int_{-\infty}^{\infty} \frac{1}{1+(\frac{x-x_0}{\exp(\gamma)})^2} dx = 1$

$$= \frac{1}{\pi \exp(\gamma)} \int_{-\infty}^{\infty} \frac{x-x_0}{1+(\frac{x-x_0}{\exp(\gamma)})^2} dx + x_0$$

$$= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\frac{x-x_0}{\exp(\gamma)}}{1+(\frac{x-x_0}{\exp(\gamma)})^2} dx$$

let $\dfrac{x-x_0}{\exp(\gamma)} = u$ $\quad du = \dfrac{1}{\exp(\gamma)} dx$ $\quad dx = du \cdot \exp(\gamma)$

$$= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{u}{1+u^2} \exp(\gamma) du + x_0$$

let $m = u^2$ $\quad dm = 2u\,du$ $\quad du = \dfrac{dm}{2u}$

$$= \frac{\exp(\gamma)}{\pi} \int_{-\infty}^{\infty} \cancel{u} \cdot \frac{1}{1+m} \cdot \frac{dm}{2\cancel{u}} = \frac{\exp(\gamma)}{\pi} (-\ln(1-x)) \Big|_{-\infty}^{\infty}$$

since $\ln(-\infty)$ is undefined. $E(X)$ is undefined.

(c)

$$L(\{x_0, \gamma\}; \{x^{(i)}\}_{i=1}^{N})$$

$$= \prod_{i=1}^{n} \frac{1}{\pi \exp(\gamma) \left[1 + \left(\frac{x_i - x_0}{\exp(\gamma)}\right)^2\right]}$$

$$= \frac{1}{N\pi \cdot N\exp(\gamma) \cdot \prod_{i=1}^{n} \left[1 + \left(\frac{x_i - x_0}{\exp(\gamma)}\right)^2\right]}$$

$$\ln L(\{x_0, \gamma\}; \{x^{(i)}\}_{i=1}^{N})$$

$$= -N\ln(\pi) - N\gamma - \sum_{i=1}^{N} \ln\left(1 + \left(\frac{x_i - x_0}{\exp(\gamma)}\right)^2\right)$$

$$\frac{d \ln L(\{x_0, \gamma\}; \{x^{(i)}\}_{i=1}^{N})}{dx_0} = -\sum_{i=1}^{n} \frac{1}{1 + \left(\frac{x_i - x_0}{\exp(\gamma)}\right)^2} \cdot (-2) \cdot \frac{x_i - x_0}{\exp(\gamma)} \cdot \frac{1}{\exp(\gamma)}$$

$$= \frac{2}{\exp(\gamma)} \sum_{i=1}^{N} \frac{\left(\frac{x - x_0}{\exp(\gamma)}\right)}{1 + \left(\frac{x - x_0}{\exp(\gamma)}\right)^2}$$

$$\frac{d \ln L(\{x_0, \gamma\}; \{x^{(i)}\}_{i=1}^{N})}{d\gamma} = -\sum_{i=1}^{N} \frac{1}{1 + \left(\frac{x_i - x_0}{\exp(\gamma)}\right)^2} \cdot 2 \cdot \frac{x_i - x_0}{\exp(\gamma)} \cdot \left(-\frac{x_i - x_0}{\exp(\gamma)}\right) \cdot \exp(\gamma) - N$$

$$= -N + \sum_{i=1}^{N} \frac{2\left(\frac{x_i - x_0}{\exp(\gamma)}\right)^2}{1 + \left(\frac{x_i - x_0}{\exp(\gamma)}\right)^2}$$

Question 4.

a) $$\frac{dL}{db} = -\frac{1}{N}\sum_{i=1}^{N}(y^{(i)} - (b + w^T \cdot x^{(i)}))^2$$

$$\frac{dL}{dw} = \frac{1}{N}\sum_{i=1}^{N}(y^{i} - (b + w^T \cdot x^{i}))^2 + 2\lambda w$$

$$b_{new} = b - \eta \cdot \frac{dL}{db}$$

$$w_{new} = w - \eta \cdot \frac{dL}{dw}$$