1.

a. let $Z_i = -\log(U_i)$.

According to Convolution:

The density of $Z = X + Y$ is given by the convolution of f & g: $[f * g](z) := \int f(z-y)g(y)\,dy$.

for $Y := -\sum_{i=1}^{n} \log(U_i)$

$PDF(Y) = PDF(Z_1) * PDF(Z_2) * \cdots * PDF(Z_n)$.

① PDF of Z: $f(z) = f(u) \cdot \left|\frac{du}{dz}\right|$

$z = -\log u$. $\begin{cases} \frac{dz}{du} = -\frac{1}{u} \\ \frac{du}{dz} = -u. \end{cases}$

$f(u) = \frac{1}{b}$. $\left|\frac{du}{dz}\right| = u$. $\boxed{f(z) = \frac{u}{b}}$

$z = -\log u \Rightarrow u = e^{-z}$.

$\Rightarrow \boxed{f(z) = e^{-z} \frac{1}{b}}$ $\qquad e^{z-w} \frac{1}{b} \cdot e^{-z} \frac{1}{b}$.

②. PDF of $Z_1 + Z_2$:

We now introduce a new parameter: w

for the new PDF of $Z_1 + Z_2$.

$$[f*g](w) = \int f(w-z)g(z)\,dz.$$

in this case: $f = g = e^{-z}\frac{1}{b}$.

$$[f*g](w) = \int e^{z-w}\frac{1}{b}\cdot e^{-z}\frac{1}{b}\,dz$$

$$= \int e^{-w}\frac{1}{b^2}\,dz.$$

$\because$ $Z = -\log(U)$, $U \in (0, b)$.

$$\begin{cases} \text{Domain of } g(z): (-\log b, \infty). \\ \text{Domain of } f(w-z): w-z > -\log b \\ \quad \Rightarrow z < w + \log b. \end{cases}$$

$\Rightarrow$ Range of Z: $(-\log b, w+\log b)$.

Therefore: $[f*g](w) = \int_{-\log b}^{w+\log b} e^{-w}\frac{1}{b^2}\,dz.$

$$= e^{-w}\frac{1}{b^2}z\Big|_{-\log b}^{w+\log b} = \boxed{\frac{1}{b^2}e^{-w}(w+2\log b)}$$

$w > -2\log b.$

③. PDF of $Z_1+Z_2+Z_3$. $\Rightarrow$. PDF of $(Z_1+Z_2)+Z_3$.

$$\begin{cases} \text{PDF of } Z_1+Z_2 : g : \frac{1}{b^2} e^{-w}(w+2\log b) \\ \text{PDF of } Z_3 : \quad f : e^{-z}\frac{1}{b}. \end{cases}$$

now we introduce a new parameter w_2.

$$[f*g](w_2) = \int f(w_2 - z)\, g(z)\, dz.$$

$$= \int e^{z-w_2}\frac{1}{b} \cdot \frac{1}{b^2} e^{-z}(z+2\log b)\, dz.$$

$$= \int e^{-w_2}\frac{1}{b^3}(z+2\log b)\, dz.$$

Range. $\begin{cases} z \in (-\log b, +\infty). \\ w_2 - z > -\log b. \quad z < w_2 + \log b. \end{cases}$

$\Rightarrow$. $z \in (-2\log b, w_2 + \log b)$.

$\therefore$ $[f*g](w_2) = \boxed{\frac{1}{b^3} e^{-w_2} \cdot \frac{1}{2}(w_2+3\log b)^2}$

$\Rightarrow$ $P(y;b) = \boxed{\frac{e^{-y}}{(n-1)!\, b^n}(y+n\log b)^{n-1}}$.

$y \in (-n\log b, \infty)$

b. It's obvious that the PDF has the same form of Gamma distribution.

$$P(y_i; b) = \frac{e^{-y}}{(n-1)!\,b^n}(y + b\log b)^{n-1}$$

$$= \frac{e^{(-y - n\log b)}}{(n-1)!}(y + b\log b)^{n-1} = \frac{e^{-(y+n\log b)}}{\Gamma(n)}(y + b\log b)^{n-1}$$

$$= \frac{1}{\Gamma(n)\theta^n} x^{n-1} e^{-\frac{x}{\theta}}$$

$$\because \quad x = y + n\log b \quad \& \quad \theta = 1$$

$$E[Y^3] = \frac{n(n+1)(n+2)}{\theta^3} = n(n+1)(n+2)$$

$$\boxed{= n^3 + 3n^2 + 2n.}$$

c. $$P(y|b) = \frac{e^{-y}}{(n-1)!\,b^n} \cdot (y + n\log(b))^{n-1}$$

$$l(b) = \log \prod_{1}^{2} f(y_i|b)$$

$$= \sum_{i=1}^{2} \log f(y_i|b)$$

$$= \sum_{i=1}^{2} \left[-y + (n-1)\log(y + n\log(b)) - \log(b^n(n-1)!)\right]$$

$$= \sum_{i=1}^{2} [-y + (n-1)\log(y + n\log(b)) - n\log(b) - \log(n-1)!]$$

$$\frac{dl(b)}{db} = \sum_{i=1}^{2} \left(-\frac{n}{b} + \frac{n-1}{y + n\log b} \cdot \frac{n}{b}\right)$$

$$= \frac{n}{b}\left(\frac{n-1}{y_1 + n\log b} + \frac{n-1}{y_2 + n\log b} - 2\right). = A.$$

$$l'(b) = 0 \implies A = 0 \implies \frac{n-1}{y_1 + n\log b} + \frac{n-1}{y_2 + n\log b}\frac{n}{b} = 2\frac{n}{b}$$

$$\implies$$

$$\boxed{b = \exp\left(\frac{\pm\sqrt{n^2 - 2n + y_1^2 + y_2^2 + 1 - 2y_1y_2} + n - y_1 - y_2 - 1}{2n}\right)}$$

2. (a)

① $\boxed{\text{Dk.}}$ for both one & k classes.

② $\boxed{k \cdot (D + \frac{D^2}{2})}$ for both one & k-clas

③ $\boxed{2D}$ for one class. $\boxed{2DK}$ for k classes.

④ $\boxed{D+1}$ for one class. $\boxed{k(D+1)}$ for k classes.

b.

$$J(w_1 \sim w_k) = \log \prod_{i=1}^{n} e_{y_i}^T \frac{\exp(w_k^T x_i)}{\sum_{j=1}^{k} \exp(w_j^T x_i)} = \sum_{i=1}^{n} e_{y_i}^T \log\left(\frac{e^{w_k^T x_i}}{\sum_1^k e^{w_k^T x_i}}\right)$$

$$\frac{dJ}{dw_{d,k}} = \sum_{j=1}^{k} \frac{dJ}{ds(w_j^T x_i)} \frac{ds(w_j^T x_i)}{d(w_k^T x_i)} \frac{d(w_k^T x_i)}{d(w_{d,k})}$$

$= x_{i,d}$.

$$\frac{\exp(w_k^T x_i)}{\sum_{j=1}^{k} \exp(w_j^T x_i)} \Rightarrow \text{a softmax function.}$$

$$s(w_j^k x_i) - s(w_j^T x_i)\, s(w_j^T x_i)$$

$$\frac{e_{y_i}^T}{s(w_j^T x_i)}$$

Therefor: $$\frac{dJ}{dw_{d,k}} = \sum_{j=1}^{k} \frac{e_{y_i}^T}{s(w_j^T x_i)} \left(s(w_j^T x_i)_{d_{j=k}} - s(w_j^T x_i)\right) s(w_k^T x_i) \cdot x_{i,d}.$$

$$= (e_{y_k}^T - s(w_k^T x_i))\, x_{d,i} + \underline{\underline{0}}$$

Matrix form: $$\frac{dJ}{dw_k} = \sum_{i=1}^{n} (e_{y_k}^T - s(w_k^T x_i))\, x_i^T$$

$$\boxed{w_k^{t+1} = w_k^t + \alpha \sum_{i=1}^{n} (e_{y_k}^T - s(w_k^T x_i))\, x_i^T}$$

3.

a. $D = \left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ x_{42} \end{pmatrix}, \begin{pmatrix} x_{51} \\ 6 \end{pmatrix} \right\}$

$$Q(\theta, \theta^{(0)}) = \int \log P(D|\theta)\, P_z(z|D, \theta)\, dz.$$

$$= \iint \left[\sum_{i=1}^{5} \log P(x_i|\theta) \right] P(x_{42}|D, \theta^{(0)})\, P(x_{51}|D, \theta^{(0)})\, dx_{51}\, dx_{42}.$$

$$= \sum_{i=1}^{3} \log P(x_i|\theta) + \int \log P(x_4|\theta) P(x_{42}|x_4, \theta^{(0)})\, dx_{42} + \int \log P(x_5|\theta) P(x_{51}|x_5, \theta^{(0)})\, dx_{51}$$

$$P(x_{ij}|x_i, \theta^{(0)}) = \frac{P(x_4|\theta^{(0)})}{\int P(x_4|\theta^{(0)})\, dx} = 10 P(x_i|\theta^{(0)}).$$

$$P(x_{42}|\theta^{(0)}) = \begin{cases} 1/10 & 0 \le x_{42} \le 10 \quad (l_2^{(0)}, u_2^{(0)}). \\ 0 & \text{otherwise.} \end{cases}$$

$$P(x_{51}|\theta^{(0)}) = \begin{cases} 1/10 & 0 \le x_{51} \le 10 \quad (l_1^{(0)}, u_2^{(0)}) \\ 0 & \text{otherwise.} \end{cases}$$

$$Q(\theta, \theta^{(0)}) = \boxed{-10 \log 10}$$

b.

$$T(\theta,\theta^{(i)}) = -3\{\log(u_1-l_1)+\log(u_2-l_2)$$

$$-\int_{l_1}^{u_1}\left[\log(u_1-l_1)+\log(u_2-l_2)\right]\frac{1}{u_1-l_1}\,dx_{51}$$

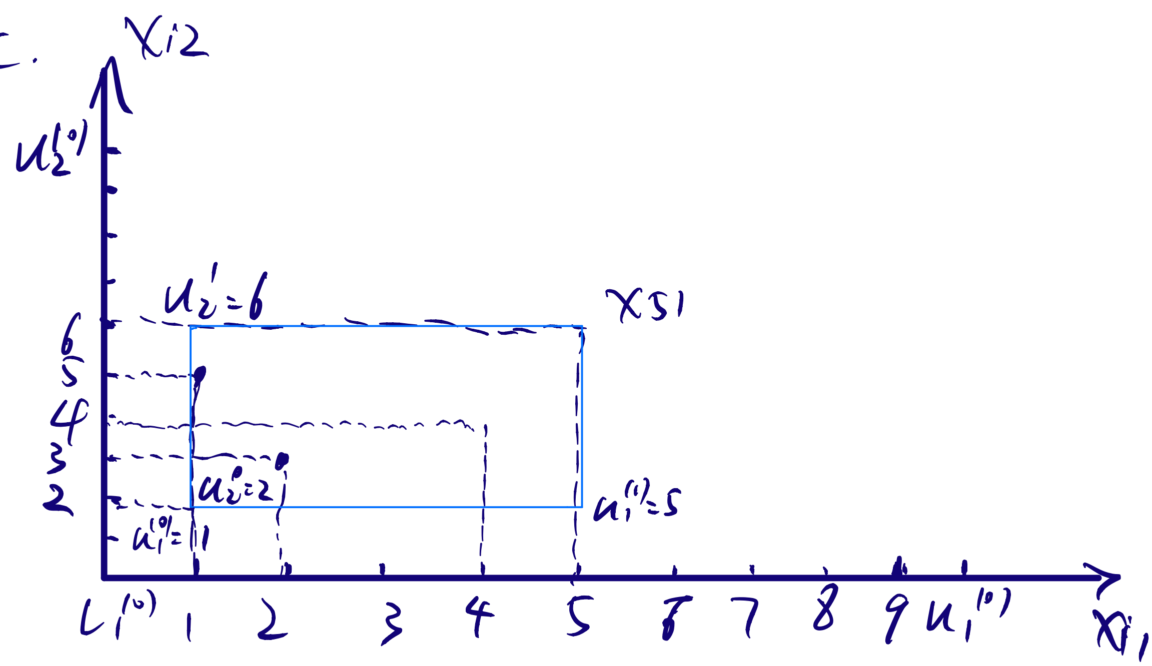
$$-\int_{l_2}^{u_2}\left[\log(u_1-l_1)+\log(u_2-l_2)\right]\frac{1}{u_2-l_2}\,dx_{42}.$$

$$= -\left[\log(u_1-l_1)+\log(u_2-l_2)\right]\cdot\Big[3+\underbrace{\int_{l_1}^{u_1}\frac{1}{u_1-l_1}\,dx_{51}}_{A}+\underbrace{\int_{u_2}^{l_2}\frac{1}{u_2-l_2}\,dx_{42}}_{B}\Big]$$

$\theta^{(1)} = [1,5,2,6]$

for $A, B. = \begin{cases} 0 & \text{outside} \\ 1 & \text{inside} \end{cases}$

c.



d.

$$\max_{\theta MLE} (D_{known}, \theta MLE)$$

$$= \max_{\theta MLE} P(x_1|\theta) P(x_2|\theta) P(x_3|\theta)$$

$$= \max_{\theta MLE} \sum_{i=1}^{3} \log P(x_i|\theta)$$

$\theta MLE = [1, 4, 2, 4]$. → smallest box that contains x_1, x_2, x_3

$P = 3$ is the size of the box.