1.1

1

$$P(x \mid p) = \prod_{d=1}^{D} p_d^{x_d} (1 - p_d)^{1 - x_d}.$$

2

$$P(x^{(i)} \mid \mathbf{p}, \pi) = \sum_{k=1}^{K} P(x^{(i)} \mid p^{(k)}) \pi(k).$$

3

The log likelihood is

$$\sum_{i=1}^{n} \log P(x^{(i)} \mid \mathbf{p}, \pi).$$

1.2

1

The probability that $p^{(k)}$ was chosen for $x^{(i)}$ to be drawn from is $\pi(k)$. So,

$$P(z^{(i)} \mid \pi) = \sum_{k=1}^{K} \pi(k) z_k^{(i)}.$$

Edit: This can also be written as

$$\prod_{k=1}^{K} \pi(k)^{z_k^{(i)}}.$$

For the second question, the indicator variable tells us that the distribution we are drawing from is

$$p = \sum_{k=1}^{K} z_k^{(i)} p^{(k)} = \langle z^{(i)}, \mathbf{p} \rangle.$$

So, the probability $x^{(i)}$ is selected is

$$P\left(x^{(i)} \mid \langle z^{(i)}, \mathbf{p} \rangle\right)$$

Note that π isn't actually a factor. Edit: This can be written another way as

$$\prod_{k=1}^{K} \left(\prod_{d=1}^{D} p_d^{x_d^{(i)}} (1 - p_d)^{1 - x_d^{(i)}} \right)^{z_k}.$$

2

By Bayes' law (and assuming independence of the $x^{(i)}$),

$$\begin{aligned} P(X, Z \mid \pi, \mathbf{p}) &= P(X \mid Z, \pi, \mathbf{p}) P(Z \mid \pi, \mathbf{p}) \\ &= P(X \mid Z, \pi, \mathbf{p}) P(Z \mid \pi) \\ &= \left(\prod_{i=1}^{n} P(x^{(i)} \mid z^{(i)}, \pi, \mathbf{p}) \right) \left(\prod_{i=1}^{n} P(z^{(i)} \mid \pi) \right). \end{aligned}$$

3

Since $z^{(i)}$ is an indicator variable,

$$\eta\left(z_k^{(i)}\right) = E\left[z_k^{(i)} \mid x^{(i)}, \pi, \mathbf{p}\right] = P\left(z_k^{(i)} = 1 \mid x^{(i)}, \pi, \mathbf{p}\right).$$

By Bayes' theorem,

$$\begin{aligned} P\left(z_k^{(i)} = 1 \mid x^{(i)}, \pi, \mathbf{p}\right) &= \frac{P\left(x^{(i)} \mid z_k^{(i)} = 1, \pi, \mathbf{p}\right) P(z_k^{(i)} = 1 \mid \pi, \mathbf{p})}{P\left(x^{(i)} \mid \pi, \mathbf{p}\right)} \\ &= \frac{P\left(x^{(i)} \mid p^{(k)}\right) \pi_k}{P\left(x^{(i)} \mid \pi, \mathbf{p}\right)} \end{aligned}.$$

All of these are answers we already have. Plugging those in, we get

$$\eta\left(z_k^{(i)}\right) = \frac{\pi_k \prod_{d=1}^{D} \left(p_d^{(k)}\right)^{x_d^{(i)}} \left(1 - p_d^{(k)}\right)^{1-x_d^{(i)}}}{\sum_{j=1}^{n} \pi_j \prod_{d=1}^{D} \left(p_d^{(j)}\right)^{x_d^{(i)}} \left(1 - p_d^{(j)}\right)^{1-x_d^{(i)}}},$$

as desired.

For the second part, note that we have already calculated $P(X, Z \mid \tilde{\pi}, \tilde{\mathbf{p}})$, so we can just write down

$$\begin{aligned}
\log P(X, Z \mid \tilde{\pi}, \tilde{\mathbf{p}}) &= \log \left(\left(\prod_{i=1}^{n} P(x^{(i)} \mid z^{(i)}, \tilde{\pi}, \tilde{\mathbf{p}}) \right) \left(\prod_{i=1}^{n} P(z^{(i)} \mid \tilde{\pi}) \right) \right) \\
&= \sum_{i=1}^{n} \log \left(\prod_{k=1}^{K} \left(\prod_{d=1}^{D} \tilde{p}_d^{(k)})^{x_d^{(i)}} (1 - \tilde{p}_d^{(k)})^{1-x_d^{(i)}} \right)^{z_k^{(i)}} \right) \\
&\quad + \sum_{i=1}^{n} \log \left(\prod_{k=1}^{K} \tilde{\pi}(k)^{z_k^{(i)}} \right) \\
&= \sum_{i=1}^{n} \sum_{k=1}^{K} z_k^{(i)} \left[\log \tilde{\pi}_k + \sum_{d=1}^{D} x_d^{(i)} \log \tilde{p}_d^{(k)} + (1 - x_d^{(i)}) \log(1 - \tilde{p}_d^{(k)}) \right]
\end{aligned}.$$

Taking the conditional expectation of this gives the desired result (since $E(z_i \mid X, \pi, \mathbf{p}) = \eta(z_k^{(i)})$ and the other values are constant).

1.3

1

The partial derivative with respect to $\tilde{p}_d^{(k)}$ of the E step is

$$\sum_{i=1}^{N} \eta(z_k^{(i)}) \left(\frac{x_d^{(i)}}{\tilde{p}_d^{(k)}} - \frac{1 - x_d^{(i)}}{1 - \tilde{p}_d^{(k)}} \right).$$

To find the maximum, we want this to equal zero. Multiply everything by $\tilde{p}_d^{(k)}(1 - \tilde{p}_d^{(k)})$ and collect terms to get that

$$\sum_{i=1}^{N} \eta(z_k^{(i)})(x_d^{(i)} - \tilde{p}_d^{(k)}) = 0.$$

Solving for $\tilde{p}_d^{(k)}$, we get that

$$\tilde{p}_d^{(k)} = \frac{\sum_{i=1}^N \eta(z_k^{(i)}) x_d^{(i)}}{N_k},$$

as desired.

2

We want to maximize

$$\sum_{i=1}^N \sum_{k=1}^K \eta(z_k^{(i)}) \tilde{\pi}_k = \sum_{k=1}^K N_k \tilde{\pi}_k.$$

subject to $\sum_{k'} \tilde{\pi}_{k'} = 1$. Using Lagrange multipliers, the above is maximized when

$$\langle N_1, N_2, \ldots, N_K \rangle = \lambda \langle 1, 1, \ldots, 1 \rangle$$

for some constant λ. This allows us to easily see that

$$\tilde{\pi}_k = \frac{N_k}{\sum_{k'} N_k},$$

as desired.