## 3(c) Hint:

**Problem:**

Objective function:

$$l(\theta_{0,1}, \dots, \theta_{0,15}, \theta_{1,1}, \dots, \theta_{1,15}) = \sum_{i=1}^{7} \sum_{k=1}^{15} x_k^{(i)} \log \theta_{y^i,k}$$

Where we know that first 3 messages are spam, hence we know their class value

$$y_1 = y_2 = y_3 = \_\_\_\_\_\_$$

We know that last 4 digits are not spam, hence we know their class value

$$y_4 = y_5 = y_6 = y_7 = \_\_\_\_\_\_$$

You can use this to rewrite objective function (by splitting outer sum) so that instead of $\theta_{y^i,k}$ you have either $\theta_{0,k}$ or $\theta_{1,k}$.

Constraints:

$$\sum_{k=1}^{15} \theta_{0,k} = 1$$

$$\sum_{k=1}^{15} \theta_{0,k} = 1$$

**Lagrangian multipliers**: If we are optimizing objective function $f(x)$ over constraints $g_j(x) = 0, j = 1, \dots n$, Lagrangian dual objective function is $f(x) + \sum_{j=1}^{n} \lambda_j g_j(x)$.

**Steps:**

1. Use above to define objective function with Lagrangian multipliers.
2. Find derivative with respect to $\theta_{0,k}$ where $k$ is fixed value from 1 to 15 ($i$ is not fixed though), and set it to zero.
3. Solve obtained equation for $\theta_{0,k}$ – you will obtain expression in terms of some $x_k^{(i)}$ (which are known values) and one of the multipliers (which is still unknown). If we only knew the value of the multiplier, we could use this expression to find all 15 values $\theta_{0,k}$.
4. To find the multiplier, we use constraint: $\sum_{k=1}^{15} \theta_{0,k} = 1$ i.e. sum of all 15 expressions found in Step 3. is equal to 1. This should now be equation containing only values $x_k^{(i)}$ (which, again, are given) and multiplier, hence it is equation in one variable – you can solve it to obtain value of the multiplier (expressed in terms of $x_k^{(i)}$).
5. Now since we know all $x_k^{(i)}$ and multiplier we can use expression from Step 3. to express solution for $\theta_{0,k}$ only in terms of $x_k^{(i)}$ and finally to find $\theta_{0,1}$ and $\theta_{0.7}$.
6. Repeat Steps 2. – 5. for $\theta_{1,k}$.

**3(d) Hint:** For finding $p(x)$ use Law of total probability.