2.60 Since there are n_i data points belonging to region i, all with probability density h_i, the likelihood function is given by:

$$p(\mathbf{X}|\mathbf{h}) = \prod_{n=1}^{N} p(\mathbf{x}_n|\mathbf{h}) = \prod_{i} h_i^{n_i}$$

The log-likelihood is given by:

$$\ln p(\mathbf{X}|\mathbf{h}) = \sum_{i} n_i \ln h_i$$

We have the normalization constraint:

$$\sum_{i} h_i \, \Delta_i = 1$$

And the non-negative contraint on h_i:

$$h_i \geq 0$$

Using Lagrange multipliers to enforce both constraints, we get the following Lagrangian:

$$\mathcal{L}(\ldots h_i \ldots, \lambda, \ldots \gamma_i \ldots) = \sum_{i} n_i \ln h_i + \lambda \left(\sum_{i} h_i \, \Delta_i - 1 \right) + \sum_{i} \gamma_i h_i$$

subject to $\gamma_i \geq 0$ and $\gamma_i h_i = 0 \; \forall i$.

After applying a constraint using a Lagrange multiplier, we can take derivative w.r.t each h_i:

$$\frac{\partial \mathcal{L}(\ldots h_i \ldots, \lambda, \ldots \gamma_i \ldots)}{\partial h_i} = \frac{n_i}{h_i} + \lambda \Delta_i + \gamma_i$$

Setting this derivative to 0, we get:

$$\frac{n_i}{h_i} + \lambda \Delta_i + \gamma_i = 0$$

$$\implies \lambda h_i \, \Delta_i + \gamma_i h_i = -n_i$$

Since $\gamma_i h_i = 0$, this becomes:

$$\implies \sum_i \lambda \, h_i \, \Delta_i = -\sum_i n_i$$

$$\implies \lambda \sum_i h_i \, \Delta_i = -\sum_i n_i$$

$$\implies \lambda \, (1) = -N$$

$$\implies \lambda = -N$$

Substituting this back, we get:

$$\frac{n_i}{h_i} - N \, \Delta_i = 0$$

$$\implies \frac{n_i}{h_i} = N \, \Delta_i$$

$$\implies h_i = \frac{n_i}{N \, \Delta_i}$$