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:

$$\sum_i h_i \, \Delta_i = 1$$

Using a Lagrange multiplier to enforce the normalization constraint:

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

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

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

Setting this derivative to 0, we get:

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

$$\Longrightarrow \lambda h_i \, \Delta_i = -n_i$$

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

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

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

$$\implies \lambda = -N$$

Substituting this back, we get:

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

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

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