5.37

$$p(\mathbf{t}|\mathbf{x}) = \sum_{k=1}^{K} \pi_k(\mathbf{x}) \mathcal{N}\left(\mathbf{t}|\boldsymbol{\mu}_k(\mathbf{x}), \sigma_k^2(\mathbf{x})\mathbf{II}\right)$$

Proving 5.158,

$$\mathbb{E}[\mathbf{t}|\mathbf{x}] = \int_{\mathbf{t}} \mathbf{t} \sum_{k=1}^{K} \pi_k(\mathbf{x}) \mathcal{N}\left(\mathbf{t}|\boldsymbol{\mu}_k(\mathbf{x}), \sigma_k^2(\mathbf{x})\mathbf{II}\right)\, d\mathbf{t}$$

$$= \sum_{k=1}^{K} \pi_k(\mathbf{x}) \int_{\mathbf{t}} \mathbf{t}\, \mathcal{N}\left(\mathbf{t}|\boldsymbol{\mu}_k(\mathbf{x}), \sigma_k^2(\mathbf{x})\mathbf{II}\right)\, d\mathbf{t}$$

$$= \sum_{k=1}^{K} \pi_k(\mathbf{x})\, \boldsymbol{\mu}_k(\mathbf{x})$$

Proving 5.160,

$$s^2(\mathbf{x}) = \mathbb{E}\left[\, ||\mathbf{t} - \mathbb{E}[\mathbf{t}|\mathbf{x}]||^2 \,|\, \mathbf{x}\,\right]$$

$$= \int_{\mathbf{t}} \left|\left|\mathbf{t} - \sum_{l=1}^{K} \pi_l(\mathbf{x})\boldsymbol{\mu}_l(\mathbf{x})\right|\right|^2 \sum_{k=1}^{K} \pi_k(\mathbf{x}) \mathcal{N}\left(\mathbf{t}|\boldsymbol{\mu}_k(\mathbf{x}), \sigma_k^2(\mathbf{xI})\right)\, d\mathbf{t}$$

$$= \sum_{k=1}^{K} \pi_k(\mathbf{x}) \int_{\mathbf{t}} \left|\left|\mathbf{t} - \sum_{l=1}^{K} \pi_l(\mathbf{x})\boldsymbol{\mu}_l(\mathbf{x})\right|\right|^2 \mathcal{N}\left(\mathbf{t}|\boldsymbol{\mu}_k(\mathbf{x}), \sigma_k^2(\mathbf{xI})\right)\, d\mathbf{t}$$

$$= \sum_{k=1}^{K} \pi_k(\mathbf{x}) \int_{\mathbf{t}} \left|\left|\mathbf{t} - \boldsymbol{\mu}_k(\mathbf{x}) + \boldsymbol{\mu}_k(\mathbf{x}) - \sum_{l=1}^{K} \pi_l(\mathbf{x})\boldsymbol{\mu}_l(\mathbf{x})\right|\right|^2$$
$$\mathcal{N}\left(\mathbf{t}|\boldsymbol{\mu}_k(\mathbf{x}), \sigma_k^2(\mathbf{x})\mathbf{I}\right)\, d\mathbf{t}$$

The norm can be expanded as:

$$\left|\left|\mathbf{t}-\boldsymbol{\mu}_k(\mathbf{x})+\boldsymbol{\mu}_k(\mathbf{x})-\sum_{l=1}^{K}\pi_l(\mathbf{x})\boldsymbol{\mu}_l(\mathbf{x})\right|\right|^2$$

$$=\left(\mathbf{t}-\boldsymbol{\mu}_k(\mathbf{x})+\boldsymbol{\mu}_k(\mathbf{x})-\sum_{l=1}^{K}\pi_l(\mathbf{x})\boldsymbol{\mu}_l(\mathbf{x})\right)^T$$
$$\left(\mathbf{t}-\boldsymbol{\mu}_k(\mathbf{x})+\boldsymbol{\mu}_k(\mathbf{x})-\sum_{l=1}^{K}\pi_l(\mathbf{x})\boldsymbol{\mu}_l(\mathbf{x})\right)$$

$$=\left(\mathbf{t}-\boldsymbol{\mu}_k(\mathbf{x})\right)^T\left(\mathbf{t}-\boldsymbol{\mu}_k(\mathbf{x})\right)$$
$$+\left(\boldsymbol{\mu}_k(\mathbf{x})-\sum_{l=1}^{K}\pi_l(\mathbf{x})\boldsymbol{\mu}_l(\mathbf{x})\right)^T\left(\boldsymbol{\mu}_k(\mathbf{x})-\sum_{l=1}^{K}\pi_l(\mathbf{x})\boldsymbol{\mu}_l(\mathbf{x})\right)$$
$$-2\left(\mathbf{t}-\boldsymbol{\mu}_k(\mathbf{x})\right)^T\left(\boldsymbol{\mu}_k(\mathbf{x})-\sum_{l=1}^{K}\pi_l(\mathbf{x})\boldsymbol{\mu}_l(\mathbf{x})\right)$$

Calculating the three integrals:

First term:

$$\int_{\mathbf{t}}\left(\mathbf{t}-\boldsymbol{\mu}_k(\mathbf{x})\right)^T\left(\mathbf{t}-\boldsymbol{\mu}_k(\mathbf{x})\right)\mathcal{N}\left(\mathbf{t}|\boldsymbol{\mu}_k(\mathbf{x}),\sigma_k^2(\mathbf{x})\mathbf{I}\right)\,d\mathbf{t}$$
$$=\sum_{m=1}^{L}\left(\int_{t_m}(t_m-\mu_{km}(\mathbf{x}))^2\,\mathcal{N}_k\,dt_m\right)$$
$$=\sum_{m=1}^{L}\sigma_k^2(\mathbf{x})$$
$$=L\,\sigma_k^2(\mathbf{x})$$

Second term:

$$\int_{\mathbf{t}} \left(\boldsymbol{\mu}_k(\mathbf{x}) - \sum_{l=1}^{K} \pi_l(\mathbf{x}) \boldsymbol{\mu}_l(\mathbf{x}) \right)^T \left(\boldsymbol{\mu}_k(\mathbf{x}) - \sum_{l=1}^{K} \pi_l(\mathbf{x}) \boldsymbol{\mu}_l(\mathbf{x}) \right)$$
$$\mathcal{N}\left(\mathbf{t} | \boldsymbol{\mu}_k(\mathbf{x}), \sigma_k^2(\mathbf{x}) \mathbf{I}\right) \, d\mathbf{t}$$

$$= \int_{\mathbf{t}} \left\| \boldsymbol{\mu}_k(\mathbf{x}) - \sum_{l=1}^{K} \pi_l(\mathbf{x}) \boldsymbol{\mu}_l(\mathbf{x}) \right\|^2 \mathcal{N}\left(\mathbf{t} | \boldsymbol{\mu}_k(\mathbf{x}), \sigma_k^2(\mathbf{x}) \mathbf{I}\right) \, d\mathbf{t}$$
$$= \left\| \boldsymbol{\mu}_k(\mathbf{x}) - \sum_{l=1}^{K} \pi_l(\mathbf{x}) \boldsymbol{\mu}_l(\mathbf{x}) \right\|^2 \int_{\mathbf{t}} \mathcal{N}\left(\mathbf{t} | \boldsymbol{\mu}_k(\mathbf{x}), \sigma_k^2(\mathbf{x}) \mathbf{I}\right) \, d\mathbf{t}$$
$$= \left\| \boldsymbol{\mu}_k(\mathbf{x}) - \sum_{l=1}^{K} \pi_l(\mathbf{x}) \boldsymbol{\mu}_l(\mathbf{x}) \right\|^2$$

Third term:

$$\int_{\mathbf{t}} \left(\mathbf{t} - \boldsymbol{\mu}_k(\mathbf{x}) \right)^T \left(\boldsymbol{\mu}_k(\mathbf{x}) - \sum_{l=1}^{K} \pi_l(\mathbf{x}) \boldsymbol{\mu}_l(\mathbf{x}) \right)$$
$$\mathcal{N}\left(\mathbf{t} | \boldsymbol{\mu}_k(\mathbf{x}), \sigma_k^2(\mathbf{x}) \mathbf{I}\right) \, d\mathbf{t}$$

$$= \left(\boldsymbol{\mu}_k(\mathbf{x}) - \boldsymbol{\mu}_k(\mathbf{x}) \right)^T \left(\boldsymbol{\mu}_k(\mathbf{x}) - \sum_{l=1}^{K} \pi_l(\mathbf{x}) \boldsymbol{\mu}_l(\mathbf{x}) \right)$$
$$= 0$$

Substituting these results back, we get:

$$s^2(\mathbf{x}) = \sum_{k=1}^{K} \pi_k(\mathbf{x}) \left(L \, \sigma_k^2(\mathbf{x}) + \left\| \boldsymbol{\mu}_k(\mathbf{x}) - \sum_{l=1}^{K} \pi_l(\mathbf{x}) \boldsymbol{\mu}_l(\mathbf{x}) \right\|^2 \right)$$