3.13 The predictive distribution is given by 3.57 (here we integrate over two parameters since both are unknown and have priors associated with them):

$$p(t|\mathbf{x},\mathbf{t}) = \int_{\mathbf{w}}\int_{\beta} p(t|\mathbf{x},\mathbf{w},\beta)\, p(\mathbf{w},\beta|\mathbf{t})\, d\mathbf{w}\, d\beta$$

Here, $p(t|\mathbf{x},\mathbf{w},\beta) = \mathcal{N}(t|\mathbf{w}^T\phi(\mathbf{x}),\beta^{-1})$,

and from the result of exercise 3.12, $p(\mathbf{w}|\mathbf{t}) = \mathcal{N}(\mathbf{w}|\mathbf{m}_N,\beta^{-1}\mathbf{S}_N)\, Gam(\beta|a_N,b_N)$.

$$\Longrightarrow p(t|\mathbf{x},\mathbf{t}) = \int_{\mathbf{w}}\int_{\beta} \mathcal{N}(t|\mathbf{w}^T\phi(\mathbf{x}),\beta^{-1})\; \mathcal{N}(\mathbf{w}|\mathbf{m}_N,\beta^{-1}\mathbf{S}_N)\, Gam(\beta|a_N,b_N)\, d\mathbf{w}$$

The product of the two Gaussians gives a constant multiplied by a posterior, which can be calculated using 2.116:

$$\int_{\mathbf{w}} \mathcal{N}(t|\mathbf{w}^T\phi(\mathbf{x}),\beta^{-1})\; \mathcal{N}(\mathbf{w}|\mathbf{m}_N,\beta^{-1}\mathbf{S}_N)$$

$$= \mathcal{N}(t|\phi(\mathbf{x})^T\mathbf{m}_N,\beta^{-1} + \phi(\mathbf{x})^T\beta^{-1}\mathbf{S}_N\phi(\mathbf{x}))$$

$$= \mathcal{N}(t|\phi(\mathbf{x})^T\mathbf{m}_N,\beta^{-1}(1 + \phi(\mathbf{x})^T\mathbf{S}_N\phi(\mathbf{x})))$$

Substituting this back into the integral, we get:

$$= \int_{\beta} \mathcal{N}(t|\phi(\mathbf{x})^T\mathbf{m}_N,\beta^{-1}(1 + \phi(\mathbf{x})^T\mathbf{S}_N\phi(\mathbf{x})))\; Gam(\beta|a_N,b_N)\, d\beta$$

Let $s = (1 + \phi(\mathbf{x})^T\mathbf{S}_N\phi(\mathbf{x}))$. This gives us:

$$= \int_{\beta} \frac{1}{\sqrt{2\pi\beta^{-1}s}} \exp\left\{-\frac{(t-\phi(\mathbf{x})^T\mathbf{m}_N)^2}{2\beta^{-1}s}\right\} \frac{1}{\Gamma(a_N)} b_N^{a_N}\, \beta^{a_N-1} \exp(-b_N\beta)\, d\beta$$

$$= \frac{1}{\Gamma(a_N)} b_N^{a_N}\, \frac{1}{\sqrt{2\pi s}} \int_{\beta} \beta^{(a_N-1+1/2)} \exp\left\{-\beta\left(b_N + \frac{(t-\phi(\mathbf{x})^T\mathbf{m}_N)^2}{2s}\right)\right\}\, d\beta$$

Let $c = \left(b_N + \frac{(t-\phi(\mathbf{x})^T\mathbf{m}_N)^2}{2s} \right)$. This gives:

$$= \frac{1}{\Gamma(a_N)} b_N^{a_N} \frac{1}{\sqrt{2\pi s}} \frac{1}{c^{(a_N+1/2)}} \int_\beta (c\beta)^{(a_N-1+1/2)} \exp\{-\beta c\} \; d(c\beta)$$

Using 1.141 to evaluate this integral, we get:

$$= \frac{1}{\Gamma(a_N)} b_N^{a_N} \frac{1}{\sqrt{2\pi s}} \frac{1}{c^{(a_N+1/2)}} \Gamma(a_N + 1/2)$$

$$= \frac{b_N^{a_N}}{\Gamma(a_N)} \left(\frac{1}{2\pi}\right)^{1/2} \left(\frac{1}{s}\right)^{1/2} \left[b_N + \frac{(t-\phi(\mathbf{x})^T\mathbf{m}_N)^2}{2s} \right]^{-a_N-1/2} \Gamma(a_N + 1/2)$$

$$= \frac{b_N^{a_N}}{\Gamma(a_N)} \left(\frac{1}{2\pi}\right)^{1/2} \left(\frac{1}{s}\right)^{1/2} s^{a_N+1/2} \left[sb_N + \frac{(t-\phi(\mathbf{x})^T\mathbf{m}_N)^2}{2} \right]^{-a_N-1/2} \Gamma(a_N+1/2)$$

$$= \frac{b_N^{a_N}}{\Gamma(a_N)} \left(\frac{1}{2\pi}\right)^{1/2} s^{a_N} \left[sb_N + \frac{(t-\phi(\mathbf{x})^T\mathbf{m}_N)^2}{2} \right]^{-a_N-1/2} \Gamma(a_N + 1/2)$$

$$= \frac{(sb_N)^{a_N}}{\Gamma(a_N)} \left(\frac{1}{2\pi}\right)^{1/2} \left[sb_N + \frac{(t-\phi(\mathbf{x})^T\mathbf{m}_N)^2}{2} \right]^{-a_N-1/2} \Gamma(a_N + 1/2)$$

Comparing this to the result in 2.158, we get:

$$\mu = \phi(\mathbf{x})^T\mathbf{m}_N$$
$$a = a_N$$
$$b = sb_N$$

$$\implies \lambda = a_N/sb_N = \frac{a_N\,(1+\phi(\mathbf{x})^T\mathbf{S}_N\phi(\mathbf{x}))^{-1}}{b_N}$$
$$\implies \nu = 2a_N$$