Prediction with model with unknown parameters

What we have is the model

$$f\left(y_t|\varphi_{t-1},\Theta\right)$$

where φ_{t-1} contains values of all delayed variables in the regression vector, i.e. $\psi_t = [u_t, \varphi_{t-1}]$ and Θ is a collection of all model parameters (usually regression coefficients and noise variance).

What we need is he predictive pdf (zero-step here for lucidity)

$$f\left(y_t|y\left(t-1\right)\right)$$

Its expression by means of the model

$$f\left(y_t|y\left(t-1\right)\right) = \int_{\Theta^*} f\left(y_t,\Theta|y\left(t-1\right)\right) d\Theta =$$

$\cdots$ we have inserted and immediately integrated out Θ $\cdots$

$$= \int_{\Theta^*} f\left(y_t|\varphi_{t-1},\Theta\right) f\left(\Theta|y\left(t-1\right)\right) d\Theta$$

$\cdots$ were we used the chain rule and omitted too much delayed outputs in the model $\cdots$

What we got is a product of the model and the posterior density for the estimate of parameters (from the last step - for the actual one we do not have y_t).

A great simplification can be achieved if we move to point estimates of parameters. Let $\hat{\Theta}_{t-1} = E\left[\Theta|y\left(t-1\right)\right]$ is a point estimate of Θ. The the posterior pdf (relying on this point estimate) can be expressed by means of the Dirac function

$$f\left(\Theta|y\left(t-1\right)\right) = \delta\left(\Theta, \hat{\Theta}_{t-1}\right)$$

(all probability is concentrated in the point estimate).

Then for the predictive pdf we have

$$f\left(y_t|y\left(t-1\right)\right) = \int_{\Theta^*} f\left(y_t|\varphi_{t-1}, \Theta\right) \delta\left(\Theta, \hat{\Theta}_{t-1}\right) d\Theta =$$

$$= f\left(y_t|\varphi_{t-1}, \hat{\Theta}_{t-1}\right)$$

according to the theory of generalized functions: $\int f\left(x\right) \delta\left(x, y\right) dx = f\left(y\right).$

The interpretation is clear: Use the model and substitute point estimated for the unknown parameters.