$$\mathscr{L}_{dropout}(w) = \frac{1}{N}\sum_{i=1}^{N} Error(y_i, \hat{y}_i) + \lambda \sum_{i=1}^{L} \| w \|^2$$

$$\min_{w} \mathscr{L}_{\text{dropout}}(w)$$

$$p(w) = \mathcal{N}(0, I)$$

$$p(w \mid x, y) = \frac{p(y \mid x, w)p(w)}{p(y \mid x)}$$

$$p(y^* \mid x^*, x, y) = \int p(y^* \mid x^*, w)\ p(w \mid x, y)\ dw$$

Prior:

Posterior:

Posterior predictive:

$\max_{\nu}$ ELBO(ν)