$$\log\, p(x) = KL\Big(q_\nu \parallel p(\cdot \mid x, y)\Big) + \mathbb{E}_{q_\nu}[\log\, p(W, X, Y)] \;-\; \mathbb{E}_{q_\nu}[\log\, q_\nu(W)]$$


MODEL EVIDENCE
=
DIVERGENCE ⩾ 0
+
EVIDENCE LOWER BOUND

$\max_{\nu}$ ELBO(ν)

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