Утверждение Максимизация вариационной нижней оценки

$$\int\limits_w q(w) \log \frac{p(y,w|X,h)}{q(w)} dw$$

эквивалентна минимизации расстояния Кульбака-Лейблера между распределением $q(w) \in \mathcal{Q}$ и апостериорным распределением параметров $p(w|y,X,h)$:

$$\hat{a} = \arg\max_{q \in \mathcal{Q}} \int\limits_w q(w) \log \frac{p(y,w|X,h)}{q(w)} dw \Leftrightarrow \hat{q} = \arg\min_{q \in \mathcal{Q}} D_{KL}(q(w)||p(w|y,X,h)),$$

Доказательство

$$D_{KL}(q(w)||p(w|y,X,h)) = \int\limits_w q(w) \log \frac{q(w)}{p(w|y,X,h)} dw = -\int\limits_w q(w) \log \frac{p(w|y,X,h)}{q(w)} dw =$$

$$= -\int\limits_w q(w) \log \frac{p(w,y|X,h)}{q(w)} dw + \int\limits_w q(w) \log p(y|X,h) dw.$$

Так как

$$\int\limits_w q(w) dw = 1$$

то получаем

$$D_{KL}(q(w)||p(w|y,X,h)) = -\int\limits_w q(w) \log \frac{p(w,y|X,h)}{q(w)} dw + \log p(y|X,h)$$

Так как $\log p(y|X,h)$ не зависит от w, то

$$\hat{q} = \arg\min_{q \in \mathcal{Q}} D_{KL}(q(w)||p(w|y,X,h)) = \arg\min_{q \in \mathcal{Q}} \left(-\int\limits_w q(w) \log \frac{p(w,y|X,h)}{q(w)} dw + \log p(y|X,h) \right) =$$

$$= \arg\max_{q \in \mathcal{Q}} \int\limits_w q(w) \log \frac{p(y,w|X,h)}{q(w)} dw$$