The lower bound is defined by :

$$\mathcal{L}(\theta, \phi; X^{(i)}) = \mathbb{E}_{q_\phi(Z|X^{(i)})}\left[-\log q_\phi(Z|X^{(i)}) + \log p_\theta(X^{(i)} \cap Z)\right]$$

$$= \mathbb{E}_{q_\phi(Z|X^{(i)})}\left[\log \frac{\log p_\theta(X^{(i)} \cap Z)}{q_\phi(Z|X^{(i)})}\right]$$

Let's show that it can be written as :

$$\mathcal{L}(\theta, \phi; X^{(i)}) = -D_{KL}(q_\phi(Z|X^{(i)}) \,||\, p_\theta(Z)) + \mathbb{E}_{q_\phi(Z|X^{(i)})}\left[\log p_\theta(X^{(i)}|Z)\right]$$

<u>Demonstration :</u>

$$-D_{KL}(q_\phi(Z|X^{(i)}) \,||\, p_\theta(Z)) + \mathbb{E}_{q_\phi(Z|X^{(i)})}\left[\log p_\theta(X^{(i)}|Z)\right]$$

Recall the Kullback-Leibler divergence:

$$D_{KL}(P \,||\, Q) = \int p(z) \log \frac{p(z)}{q(z)}\, dz$$

$$= -\int q_\phi(Z|X^{(i)}) \log \frac{q_\phi(Z|X^{(i)})}{p_\theta(Z)}\, dZ + \mathbb{E}_{q_\phi(Z|X^{(i)})}\left[\log p_\theta(X^{(i)}|Z)\right]$$

$$= \int q_\phi(Z|X^{(i)}) \log \frac{p_\theta(Z)}{q_\phi(Z|X^{(i)})}\, dZ + \mathbb{E}_{q_\phi(Z|X^{(i)})}\left[\log p_\theta(X^{(i)}|Z)\right]$$

$$= \mathbb{E}_{q_\phi(Z|X^{(i)})}\left[\log \frac{\log p_\theta(Z)}{q_\phi(Z|X^{(i)})}\right] + \mathbb{E}_{q_\phi(X^{(i)})}\left[\log p_\theta(X^{(i)}|Z)\right]$$

$$= \mathbb{E}_{q_\phi(Z|X^{(i)})}\left[\log \frac{p_\theta(Z) p_\theta(X^{(i)}|Z)}{q_\phi(Z|X^{(i)})}\right]$$

Then by **Bayes' formula :**

$$= \mathbb{E}_{q_\phi(Z|X^{(i)})}\left[\log \frac{\log p_\theta(X^{(i)} \cap Z)}{q_\phi(Z|X^{(i)})}\right]$$