Variational Autoencoders
In a variational autoencoder a distribution on $x$ is modeled by

$$P_{\Theta,\Psi}(x, z) = P_{\Theta}^{\text{gen}}(z) P_{\Psi}^{\text{dec}}(x|z)$$

$$P_{\Theta,\Psi}(x) = \mathbb{E}_{z \sim P_{\Theta}^{\text{gen}}} \left[ P_{\Psi}^{\text{dec}}(x|z) \right]$$

We would like

$$\Theta^*, \Psi^* = \arg\min_{\Theta,\Psi} \mathbb{E}_{x \sim D} \left[ \log \frac{1}{P_{\Theta,\Psi}(x)} \right] = \arg\min_{\Theta,\Psi} H(D, P_{\Theta,\Psi})$$

However, computing $P_{\Theta,\Psi}(x)$ by sampling $z$ from $P_{\Theta}^{\text{gen}}$ is very inefficient as $z$ is rarely compatible with $x$. 
The Variational Lower Bound

It would be more efficient to sample $z$ from $P_{\Theta,\Psi}(z|x)$. We approximate this marginal by an encoding distribution $P^\text{enc}_\Phi(z|x)$.

$$
\Phi^*, \Theta^*, \Psi^* = \arg\min_{\Phi, \Theta, \Psi} \mathbb{E}_{x \sim D} [-\mathcal{L}(x, \Phi, \Theta, \Psi)]
$$

$$
\log \frac{1}{P_{\Theta,\Psi}(x)} \leq -\mathcal{L}(x, \Phi, \Theta, \Psi)
$$

$$
\mathcal{L}(x, \Phi, \Theta, \Psi) = \mathbb{E}_{z \sim P^\text{enc}_\Phi(\cdot|x)} [\log P_{\Theta,\Psi}(x, z)] + H(P^\text{enc}_\Phi(\cdot|x))
$$

For Gaussian distributions gradient estimation through sampling is now feasible.
The Variational Lower Bound

\[ \mathcal{L}(x, \Phi, \Theta, \Psi) = \mathbb{E}_{z \sim P_{\Phi}^{\text{enc}}(\cdot|x)} \left[ \log P_{\Theta,\Psi}(x, z) \right] + H(P_{\Phi}^{\text{enc}}(\cdot|x)) \]

\[ = \mathbb{E}_{z \sim P_{\Phi}^{\text{enc}}(\cdot|x)} \left[ \log P_{\Theta,\Psi}(x) P_{\Theta,\Psi}(z|x) \right] + H(P_{\Phi}^{\text{enc}}(\cdot|x)) \]

\[ = \mathbb{E}_{z \sim P_{\Phi}^{\text{enc}}(\cdot|x)} \left[ \log P_{\Theta,\Psi}(x) \right] + \mathbb{E}_{z \sim P_{\Phi}^{\text{enc}}(\cdot|x)} \left[ \log P_{\Theta,\Psi}(z|x) \right] + H(P_{\Phi}^{\text{enc}}(\cdot|x)) \]

\[ = \log P_{\Theta,\Psi}(x) + \mathbb{E}_{z \sim P_{\Phi}^{\text{enc}}(\cdot|x)} \left[ \log P_{\Theta,\Psi}(z|x) - \log(P_{\Phi}^{\text{enc}}(z|x)) \right] \]

\[ = \log P_{\Theta,\Psi}(x) - KL(P_{\Phi}^{\text{enc}}(\cdot|x), P_{\Theta,\Psi}(\cdot|x)) \]

\[ \leq \log P_{\Theta,\Psi}(x) \]
Consistency Theorem

\[
\Phi^*, \Theta^*, \Psi^* = \arg\min_{\Phi, \Theta, \Psi} E_{x \sim D} \left[ -\mathcal{L}(x, \Phi, \Theta, \Psi) \right]
\]

\[
= \arg\min_{\Phi, \Theta, \Psi} E_{x \sim D} \left[ \log \frac{1}{P_{\Theta, \Psi}(x)} \right] + KL(P_{\Phi}^{\text{enc}}(\cdot|x), P_{\Theta, \Psi}(\cdot|x))
\]

\[
= \arg\min_{\Phi, \Theta, \Psi} H(D, P_{\Theta, \Psi}) + KL(P_{\Phi}^{\text{enc}}(\cdot|x), P_{\Theta, \Psi}(\cdot|x))
\]

**Consistency Theorem:** If for all \( \Theta \) and \( \Psi \) there exist \( \Phi \) such that \( P_{\Phi}^{\text{enc}}(z|x) = P_{\Theta, \Psi}(z|x) \) then

\[
\Theta^*, \Psi^* = \arg\min_{\Theta, \Psi} H(D, P_{\Theta, \Psi})
\]
Sampling from Variational Autoencoders

For Gaussian distributions samples from a variational autoencoder appear blurry.

[Alec Radford]
END