/70

20

# KL-divergence

A measure of similarity between two probability distributions

$$D_{\mathrm{KL}}(Q \| P) = \int_{-\infty}^{\infty} \mathrm{d}x \, q(x) \, \log\left(\frac{q(x)}{p(x)}\right)$$

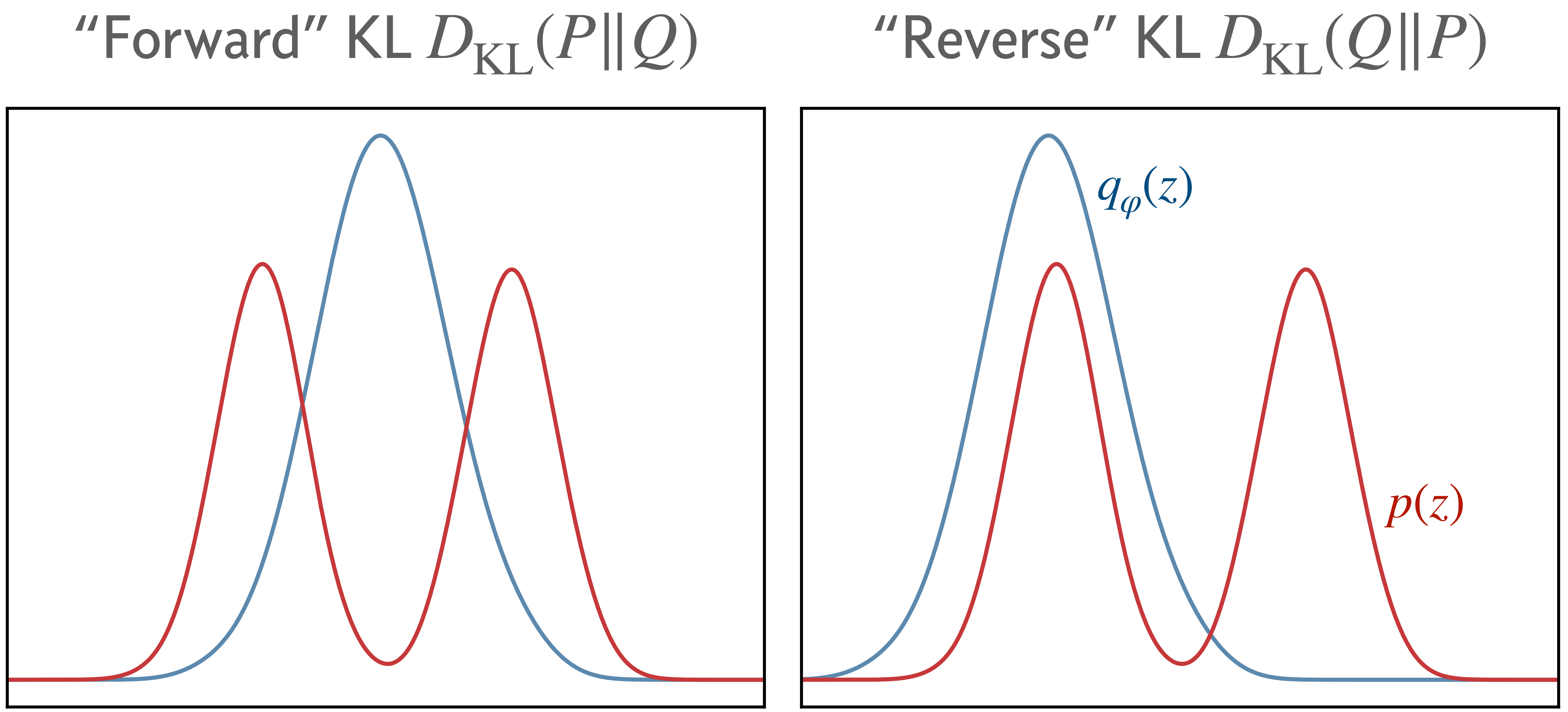
# Not symmetric!

$$D_{\mathrm{KL}}(Q \| P) \neq D_{\mathrm{KL}}(P \| Q)$$

Maximum-likelihood inference is equivalent to minimizing the *forward* KL

# Forward KL

$$D_{\mathrm{KL}}(P_{\mathscr{D}} \| Q_{\varphi}) = -\left\langle \log q_{\varphi}(z) \right\rangle_{z \sim p_{\mathscr{D}}(z)} + \mathrm{const}\,.$$


"Forward" KL $D_{\mathrm{KL}}(P\|Q)$
"Reverse" KL $D_{\mathrm{KL}}(Q\|P)$
$q_\varphi(z)$
$p(z)$

# "True" distribution

Non-negative! $D_{\mathrm{KL}}(Q \| P) \geq 0$

# KL-divergence

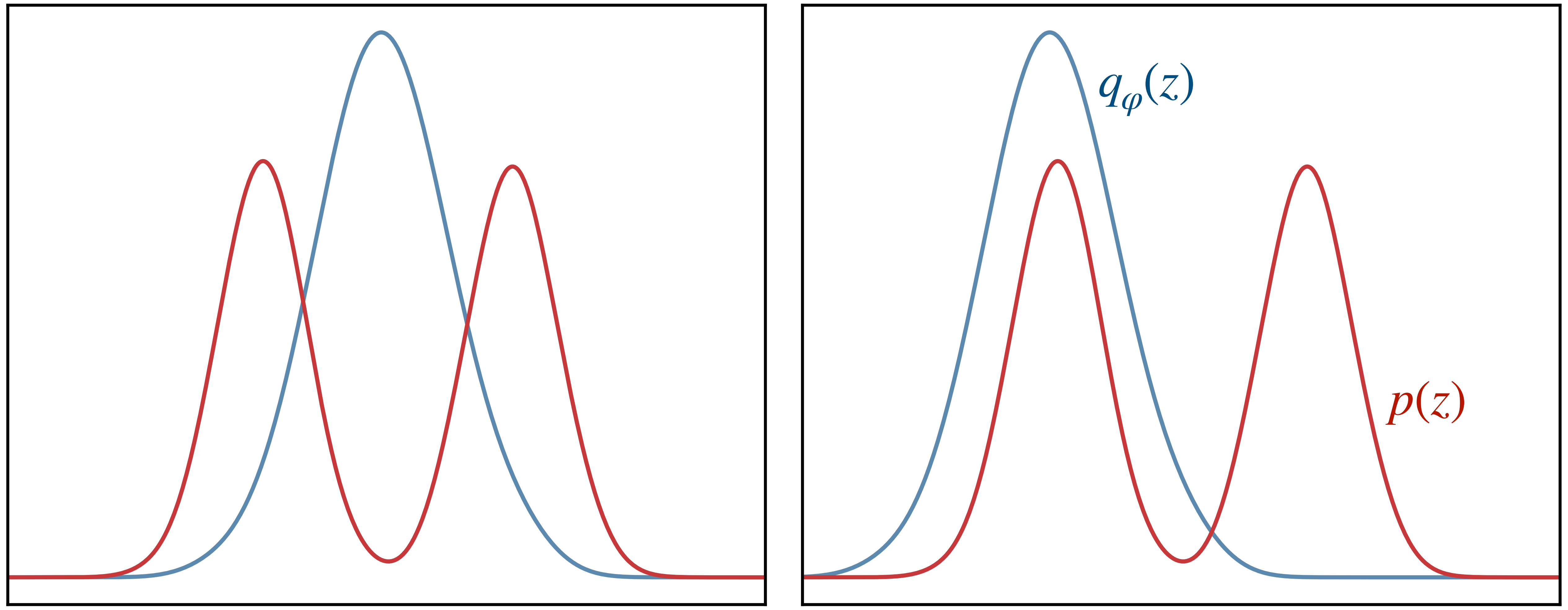
$$D_{\mathrm{KL}}(Q\|P) = \int_{-\infty}^{\infty} \mathrm{d}x\, q(x)\, \log\left(\frac{q(x)}{p(x)}\right)$$

A measure of similarity between two probability distributions

Not symmetric! $D_{\mathrm{KL}}(Q\|P) \neq D_{\mathrm{KL}}(P\|Q)$

*"True" distribution*

"Forward" KL $D_{\mathrm{KL}}(P\|Q)$ "Reverse" KL $D_{\mathrm{KL}}(Q\|P)$




Forward KL

$$D_{\mathrm{KL}}(P_{\mathscr{D}}\|Q_\varphi) = -\left\langle \log q_\varphi(z) \right\rangle_{z\sim p_{\mathscr{D}}(z)} + \text{const}\,.$$

Maximum-likelihood inference is equivalent to minimizing the *forward* KL

Non-negative! $D_{\mathrm{KL}}(Q\|P) \geq 0$

# Variational inference

Infer the posterior over the latent parameters