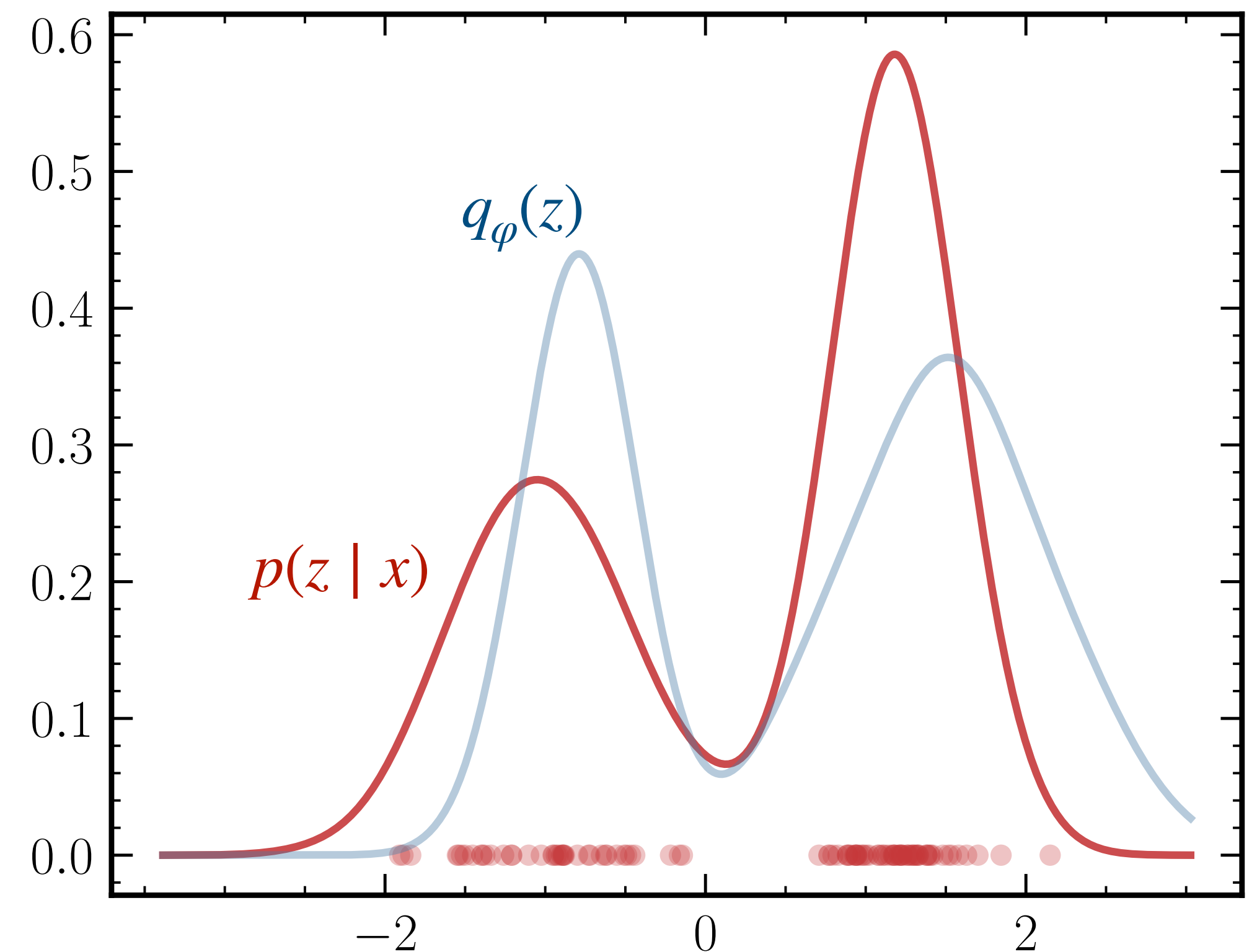


KL-divergence

A measure of similarity between two probability distributions

$$D_{\text{KL}}(Q||P) = \int_{-\infty}^{\infty} dx \, q(x) \log \left(\frac{q(x)}{p(x)} \right)$$
$$= \left\langle \log \frac{q(x)}{p(x)} \right\rangle_{x \sim q(x)}$$

Formally: *expected excess “surprise” from using P as a model when the actual distribution is Q*



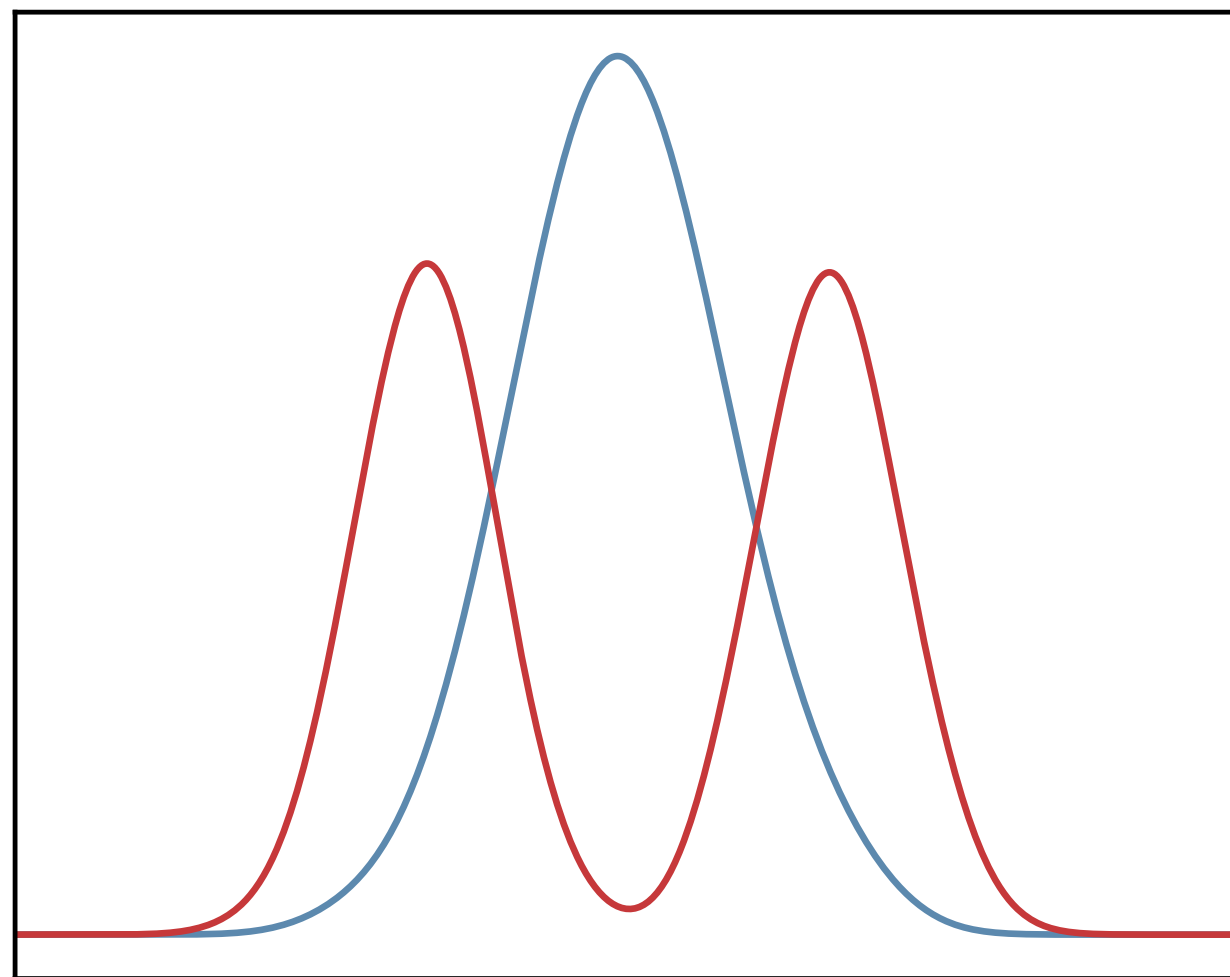
KL-divergence

A measure of similarity between two probability distributions

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

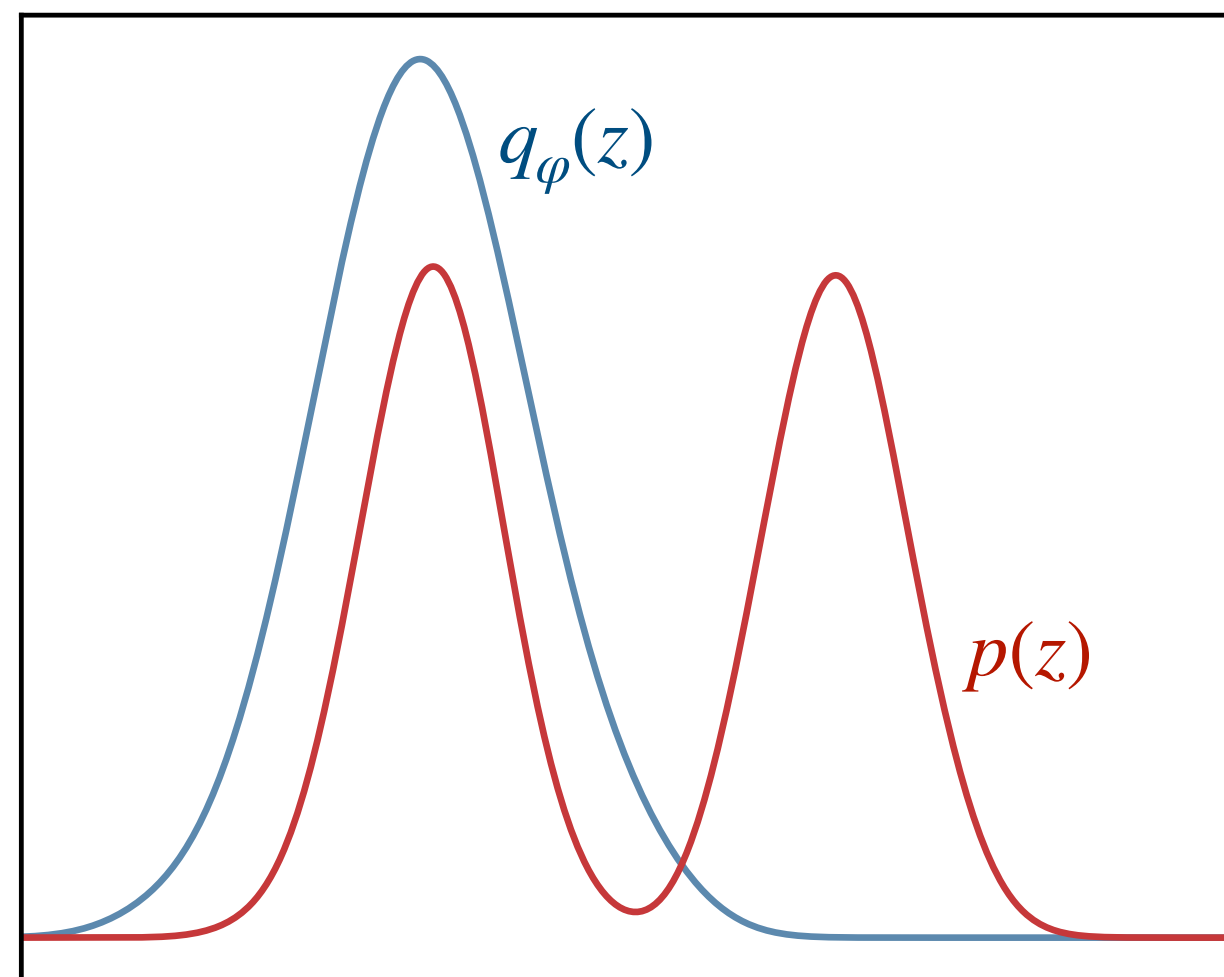
Not symmetric! $D_{\text{KL}}(Q\|P) \neq D_{\text{KL}}(P\|Q)$

“Forward” KL $D_{\text{KL}}(P\|Q)$



Mean seeking

“Reverse” KL $D_{\text{KL}}(Q\|P)$



Mode seeking