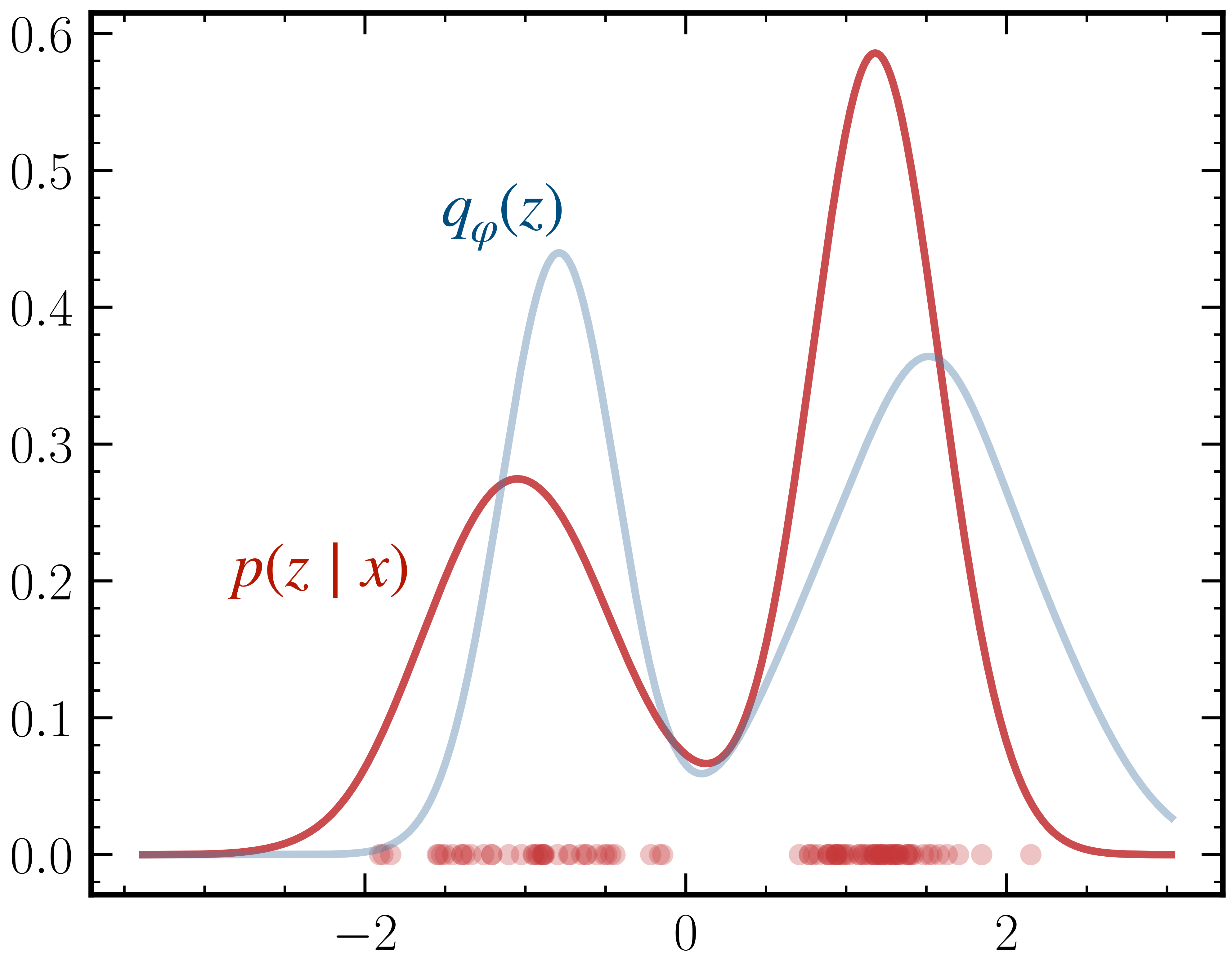
KL-divergence

A measure of similarity between two probability distributions

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

$$= \left\langle \log \frac{q(x)}{p(x)} \right\rangle_{x\sim q(x)}$$

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



KL-divergence

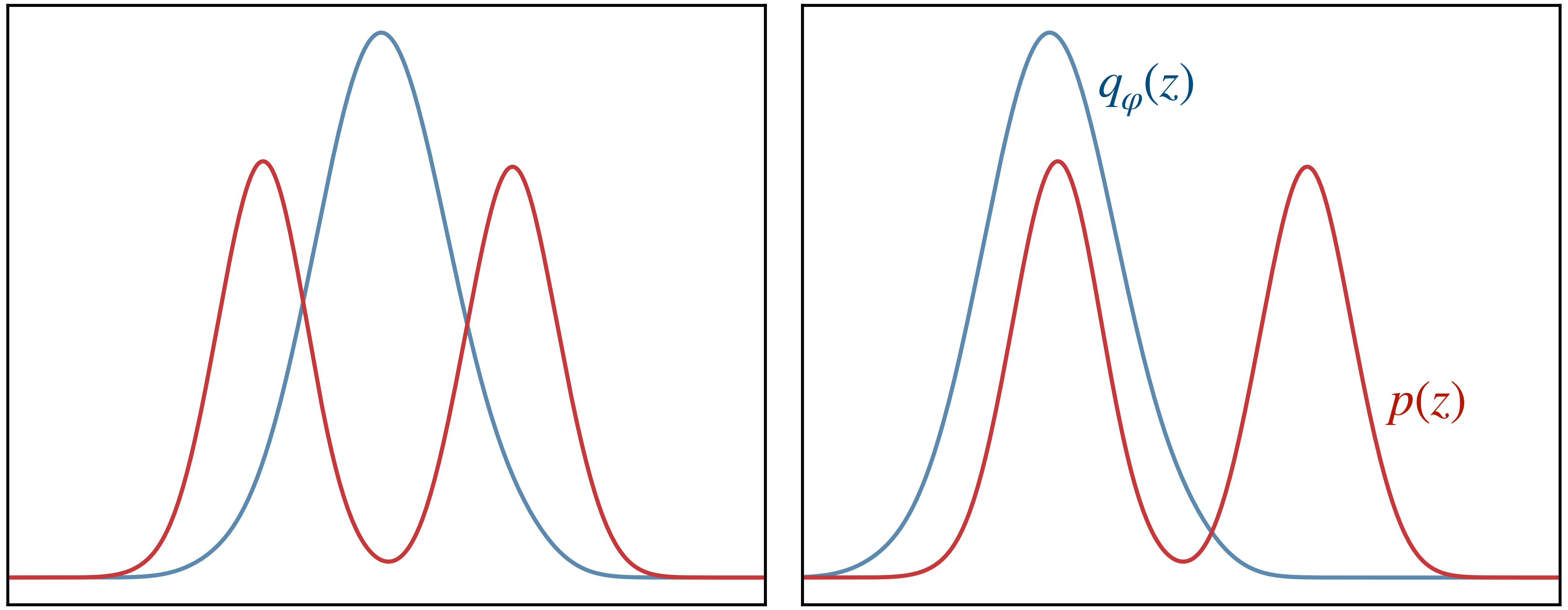
$$D_{\mathrm{KL}}(Q\|P) = \int_{-\infty}^{\infty} \mathrm{d}x\, q(x)\, \log\left(\frac{q(x)}{p(x)}\right)$$

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



Mean seeking *Mode seeking*