/70

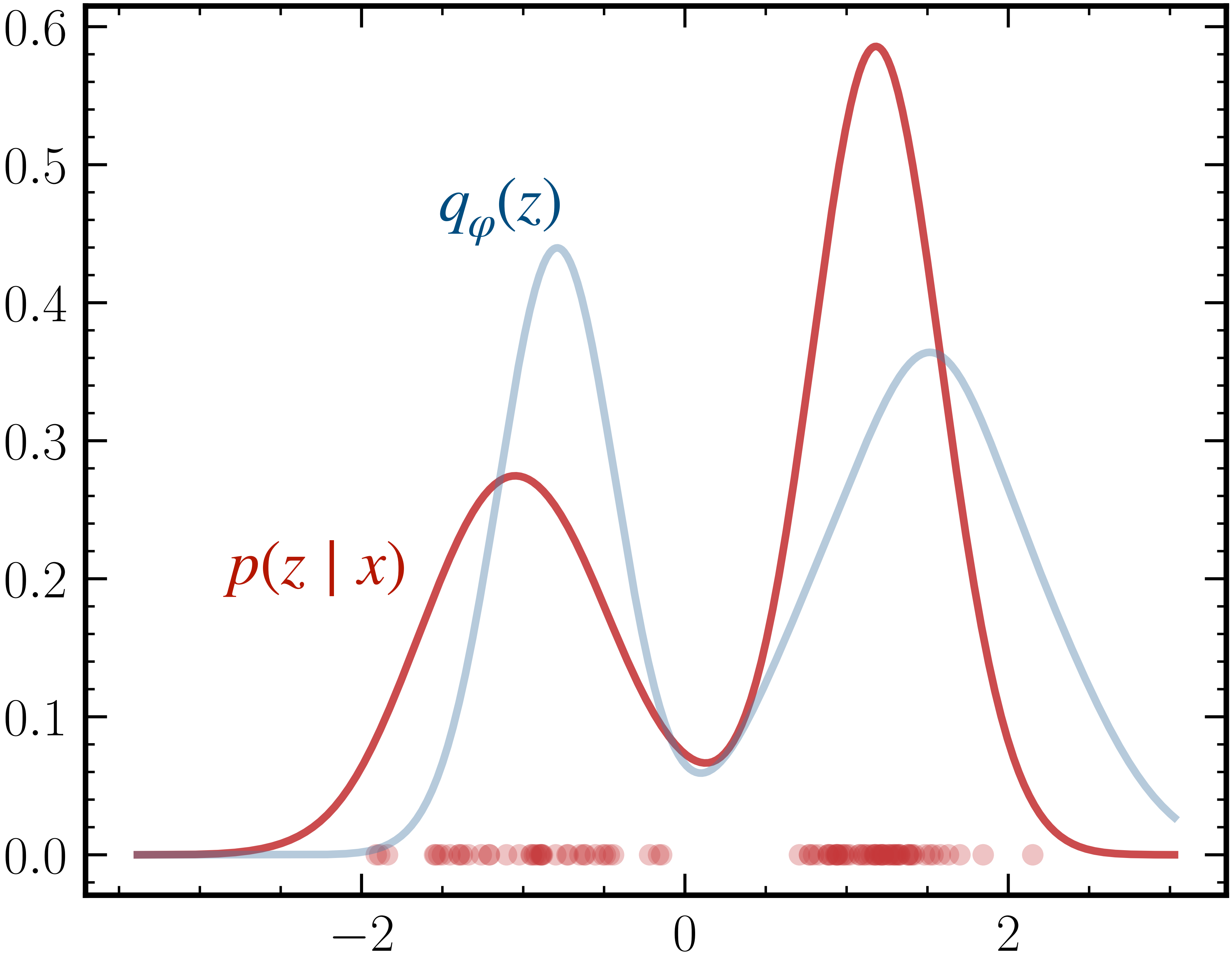
Formally: *expected excess "surprise" from using Q as a model when the actual distribution is P*

19

A measure of similarity between two probability distributions

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

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


0.6
0.5
0.4
0.3
0.2
0.1
0.0
−2
0
2
$q_\varphi(z)$
$p(z \mid x)$

Kullback–Leibler (KL) divergence

$$= -\left\langle \log q(x) \right\rangle_{p(x)} + \left\langle \log p(x) \right\rangle_{p(x)}$$

Cross-entropy $\mathbb{H}(P, Q)$ − (Self-)entropy $\mathbb{H}(P)$

Kullback–Leibler (KL) divergence

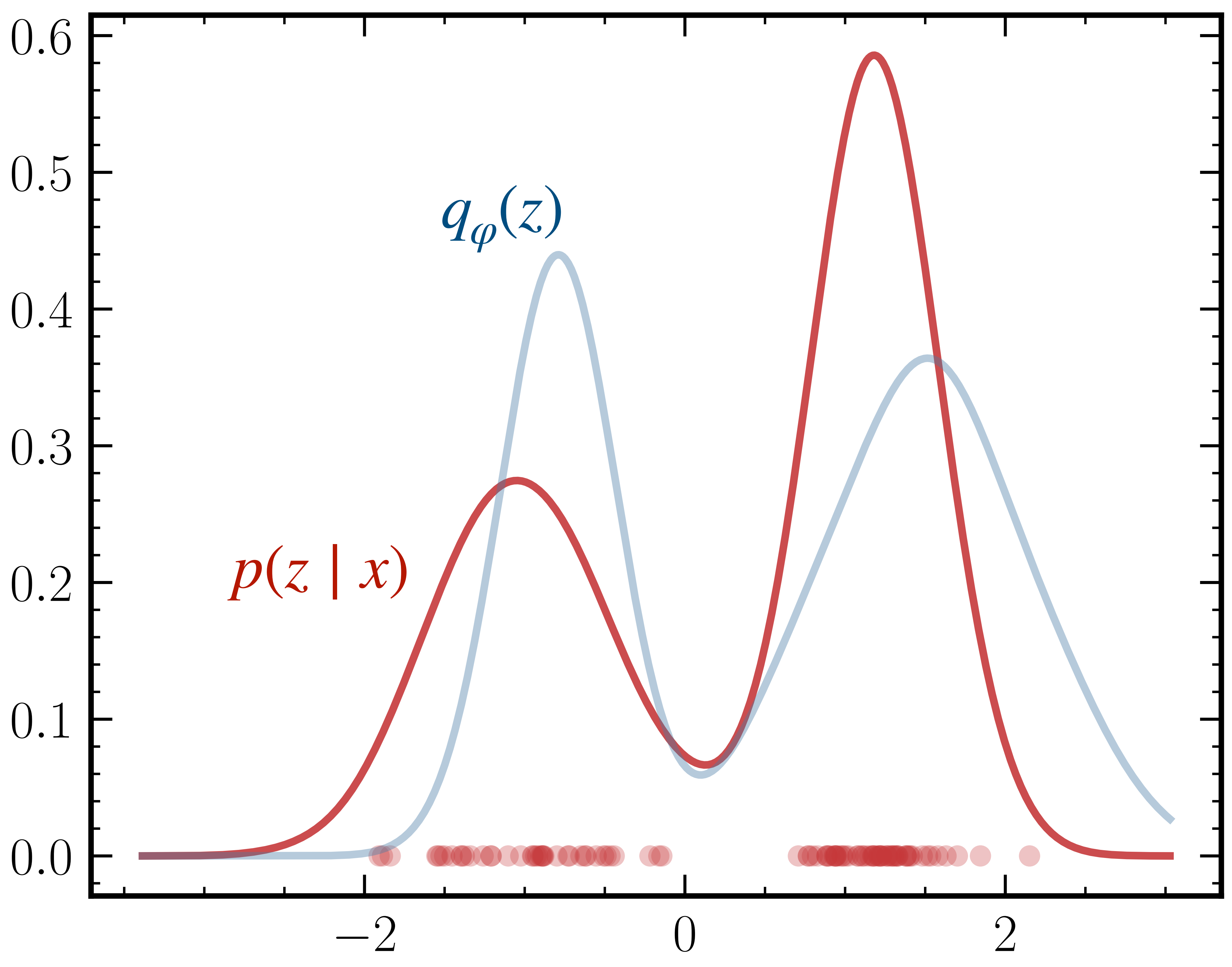
A measure of similarity between two probability distributions

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

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

$$= -\left\langle \log q(x) \right\rangle_{p(x)} + \left\langle \log p(x) \right\rangle_{p(x)}$$

Cross-entropy $\mathbb{H}(P, Q)$ − (Self-)entropy $\mathbb{H}(P)$



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

KL-divergence

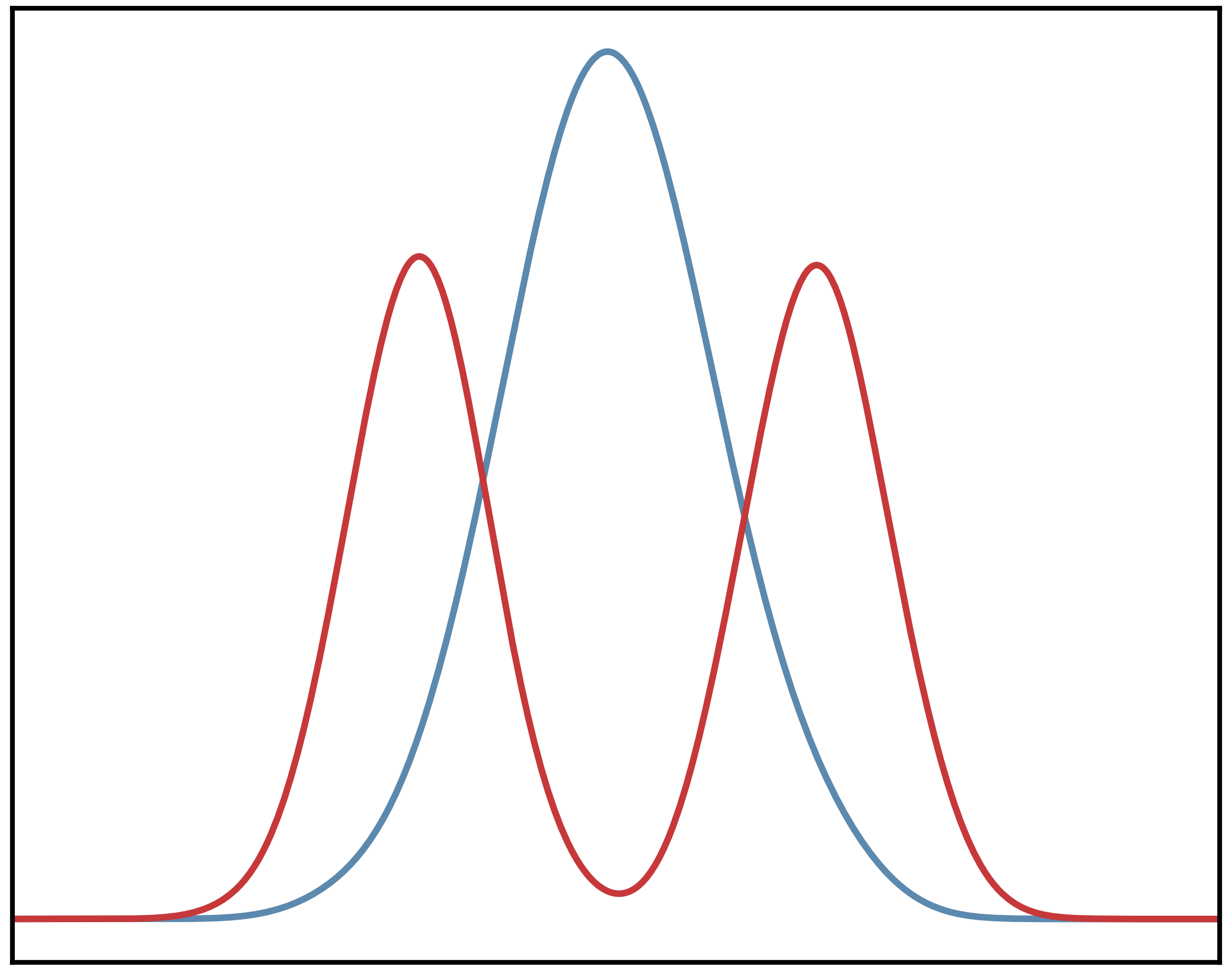
$$D_{\mathrm{KL}}(Q\|P) = \int_{-\infty}^{\infty} \mathrm{d}x\, q(x)\, \log\left(\frac{q(x)}{p(x)}\right)$$

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