Variational Inference

- Key idea: Use a **surrogate distribution** to approximate the posterior
  - Surrogate distribution is **simpler** than the **true posterior** (typically by making strong independence assumptions to break coupling)

- **Goal:** Find the "best" surrogate distribution from a certain parametric family by minimizing the KL-divergence from the surrogate (Q) to the true posterior (P)
  - Transforms inference into a deterministic optimization!

    No free lunch: approximation quality depends on variational family chosen