

- If $D(\textcolor{red}{p}||\textcolor{blue}{q})$ or $D(\textcolor{blue}{q}||\textcolor{red}{p})$ is small, then so is $V(\textcolor{red}{p}, \textcolor{blue}{q}) = V(\textcolor{blue}{q}, \textcolor{red}{p})$.
- For a sequence of probability distributions q_k , as $k \rightarrow \infty$, if $D(\textcolor{red}{p}||\textcolor{blue}{q}_k) \rightarrow 0$ or $D(\textcolor{blue}{q}_k||\textcolor{red}{p}) \rightarrow 0$, then $V(\textcolor{red}{p}, \textcolor{blue}{q}_k) = V(\textcolor{blue}{q}_k, \textcolor{red}{p}) \rightarrow 0$.
- That is, “convergence in divergence” is a stronger notion than “convergence in variational distance.”
- See Problems 23 and 24 for details.