This proof of Pinsker’s lemma is due to Pollard: \[\text{http://www.stat.yale.edu/~pollard/Books/Asymptopia/Metrics.pdf}\].

Let \(P, Q\) be two distributions over some set \(\Omega\). We’ll treat \(\Omega\) as a discrete set but all of this carries over to the continuous case as well. Denote by \(p, q\) the probability mass functions of \(P, Q\), respectively.

We have already encountered the total variation distance

\[
\frac{1}{2} \|P - Q\|_1 = \frac{1}{2} \sum_{x \in \Omega} |p(x) - q(x)|.
\]

Define the Kullback-Leibler divergence \(\text{KL}(P||Q)\):

\[
\text{KL}(P||Q) = \sum_{x \in \Omega} p(x) \log \frac{p(x)}{q(x)} = \mathbb{E}_{X \sim P} \left[ \log \frac{p(X)}{q(X)} \right].
\]

Pinsker’s inequality relates these two quantities:

**Theorem 0.1**

\[
\text{KL}(P||Q) \geq \frac{1}{2} \|P - Q\|_1^2.
\]

**Proof:** The following simple fact is left as an exercise:

\[
(1 + t) \log(1 + t) - t \geq \frac{1}{2} \frac{t^2}{1 + t/3}, \quad t \geq -1.
\]

There is no loss of generality in assuming that

\[
\sup_{x \in \Omega} \frac{p(x)}{q(x)} < \infty
\]

since otherwise \(\text{KL}(P||Q) = \infty\) and the claim is vacuously true.

Assuming (2), we may take \(p(x) = (1 + r(x))q(x)\) where \(r(x) = p(x)/q(x) - 1 \geq -1\).\(^1\)

Exercise: verify that

\[
\mathbb{E}_{X \sim Q}[r(X)] = \sum_{x \in \Omega} q(x)r(x) = 0,
\]

\[
\mathbb{E}_{X \sim Q}[|r(X)|] = \sum_{x \in \Omega} q(x)|r(x)| = \|P - Q\|_1,
\]

\[
\text{KL}(P||Q) = \mathbb{E}_{X \sim Q}[(1 + r(X)) \log(1 + r(X)) - r(X)].
\]

\(^1\)We define 0/0 = 1.
Combining (1) and (5), we get
\[
\text{KL}(P||Q) \geq \frac{1}{2} \mathbb{E}_{X \sim Q} \left[ \frac{r(X)^2}{1 + r(X)/3} \right].
\]

Now (3) implies \( \mathbb{E}_Q[1 + r(X)/3] = 1 \), and hence
\[
\text{KL}(P||Q) \geq \frac{1}{2} \mathbb{E}_Q \left[ \frac{r(X)^2}{1 + r(X)/3} \right] \mathbb{E}_Q[1 + r(X)/3].
\]

Recall the Cauchy-Schwarz inequality:
\[
(\mathbb{E}_Q[f(X)g(X)])^2 \leq \mathbb{E}_Q[f(X)^2] \mathbb{E}_Q[g(X)^2].
\]

Taking \( f(x) = \sqrt{r(x)^2/(1 + r(x)/3)} \) and \( g(x) = \sqrt{1 + r(x)/3} \), we have
\[
\text{KL}(P||Q) \geq \frac{1}{2} \mathbb{E}_Q[f(X)^2] \mathbb{E}_Q[g(X)^2] \\
\geq \frac{1}{2} \left( \mathbb{E}_Q[f(X)g(X)] \right)^2 \\
= \frac{1}{2} \mathbb{E}_Q \left[ \left( \frac{|r(X)|}{\sqrt{1 + r(X)^2}} \cdot \sqrt{1 + r(X)^2} \right)^2 \right] \\
= \frac{1}{2} \mathbb{E}_Q |r(X)|^2 \\
= \frac{1}{2} \|P - Q\|_1^2,
\]

where the last equality follows from (4).