Problem Set 2

Information and Coding Theory

February 12, 2021

Clayton Seitz

Problem 0.1. *Find tight upper and lower bounds on two extremely biased coins where the first coin is distributed according to*

$$P = \begin{cases} 0 & \epsilon \\ 1 & 1-\epsilon \end{cases}$$

and the second is distributed according to

$$Q = \begin{cases} 0 & 2\epsilon \\ 1 & 1-2\epsilon \end{cases}$$

Solution. I will assume that distinguishing the two coins means that, given a sequence of n flips, we can say whether it is coin P or coin Q in at least $\frac{9}{10}n$ trials, on average. To start, we write out the KL-Divergence between the distributions P and Q

$$\begin{aligned}
D(P||Q) &= \epsilon \log \frac{\epsilon}{2\epsilon} + (1-\epsilon)\log\frac{1-\epsilon}{1-2\epsilon} \\
&= (1-\epsilon)\log\frac{1-\epsilon}{1-2\epsilon} - \epsilon \\
&= (1-\epsilon)\log\left(1+\frac{\epsilon}{1-2\epsilon}\right) - \epsilon \\
&\leq (1-\epsilon)\left(\frac{\epsilon}{1-2\epsilon}\right) - \epsilon \\
&= \frac{1}{2\ln 2}\cdot\frac{\epsilon^2}{1-2\epsilon}
\end{aligned}$$

At the same time, we know that

$$n \geq \frac{1}{2 \ln 2 \cdot D(P||Q)} \left(\frac{8}{5}\right)^2$$

given the constraint on successful predictions. Substituting for $D(P||Q)$ gives us the lower bound on n

$$n \geq \left(\frac{1}{\epsilon^2} - \frac{2}{\epsilon}\right) \left(\frac{8}{5}\right)^2$$

$\blacksquare$

Problem 0.2. *Show that* $0 \leq \mathbf{JSD}(P, Q) \leq 1$

Solution.

$$\mathbf{JSD}(P, Q) \;=\; \frac{1}{2} D(P||M) + \frac{1}{2} D(Q||M)$$

The lower bound must be true because $D(P||M) \geq 0$ and $Q(P||M) \geq 0$. For the upper bound, consider just one of the terms

$$\begin{aligned} D(P||M) &= \frac{1}{2} \sum_{x \sim P} P(x) \log \frac{P(x)}{M(x)} \\ &= \frac{1}{2} \sum_{x \sim P} P(x) \log \frac{2P(x)}{P(x) + Q(x)} \\ &\leq \frac{1}{2} \sum_{x \sim P} P(x) \log 2 = \frac{1}{2} \end{aligned}$$

Therefore, $\mathbf{JSD}(P, Q) \leq 1$.

Show that $\mathbf{JSD}(P, Q) \geq \frac{1}{8 \ln 2} \cdot ||P - Q||_1^2$

$$\begin{aligned}
\mathbf{JSD}(P,Q) &= \frac{1}{2}\left[D(P||M)+D(Q||M)\right] \\
&\geq \frac{1}{4\ln 2}\left[||P-M||_1^2+||Q-M||_1^2\right] \\
&= \frac{1}{4\ln 2}\left[\left(\sum|P-M|\right)^2+\left(\sum|Q-M|\right)^2\right] \\
&= \frac{1}{8\ln 2}\left[\left(\sum|P-Q|\right)^2+\left(\sum|Q-P|\right)^2\right] \\
&= \frac{1}{8\ln 2}\cdot||P-Q||_1^2
\end{aligned}$$

Show that $\mathbf{JSD}(\mathbf{P},\mathbf{Q}) \leq \frac{1}{2}\cdot||P-Q||_1$

$$\begin{aligned}
\mathbf{JSD}_\lambda(P,Q) &= \frac{1}{2}\left[D(P||M)+D(Q||M)\right] \\
&= \frac{1}{2}\left[\sum_{x\sim P} P(x)\log\frac{P(x)}{M(x)}+\sum_{x\sim Q} Q(x)\log\frac{Q(x)}{M(x)}\right] \\
&= \frac{1}{2}\sum_{x\sim\chi}\left[P(x)\log\frac{2P(x)}{P(x)+Q(x)}+Q(x)\log\frac{2Q(x)}{P(x)+Q(x)}\right] \\
&= \frac{1}{2}\sum_{x\sim\chi}[P(x)+Q(x)]\cdot\frac{P(x)}{P(x)+Q(x)}\log\frac{2P(x)}{P(x)+Q(x)} \\
&\quad+\frac{Q(x)}{P(x)+Q(x)}\log\frac{2Q(x)}{P(x)+Q(x)} \\
&= \frac{1}{2}\sum_{x\sim\chi}[P(x)+Q(x)]\cdot|1-H\left(\frac{P(x)}{P(x)+Q(x)},\frac{Q(x)}{P(x)+Q(x)}\right)| \\
&\leq \sum_{x\sim\chi}|P(x)-Q(x)|
\end{aligned}$$

$$\mathbf{JSD}_\lambda(P_1\ ...\ P_k)=\sum_i \lambda_i D(P_i||M)$$

where $M=\sum_i \lambda_i P_i$. Show that

$$0 \leq \mathbf{JSD}_\lambda(P_1\ ...\ P_k) \leq H(\lambda)$$

As before, the lower bound must be true because $D(P_i||M) \geq 0$ and λ is non-negative. As for the upper bound,

$$\begin{aligned}
\mathbf{JSD}_\lambda(P_1 \ldots P_k) &= \sum_i \lambda_i D(P_i||M) \\
&= \sum_i \lambda_i P_i \log \frac{P_i}{M} \\
&= H(\sum_i \lambda_i P_i) - \sum_i \lambda_i H(P_i) \\
&= H(\lambda) - \sum_i \lambda_i H(P_i) \\
&\leq H(\lambda)
\end{aligned}$$

■

Problem 0.3. *Counting using the method of types*

Solution. Sanov's Theorem states that

$$Q^n(E) = (n+1)^r 2^{-nD(P^*||Q)}$$

where P^* is the distribution in E that is closest to Q. Since Q is a uniform distribution we have that

$$D(P^*||Q) = \log m - H(P)$$

Therefore, if we let H^* be the entropy of the distribution which has maximum entropy, Sanov's theorem becomes

$$Q^n(E) = (n+1)^r 2^{-n(\log m - H^*)}$$

$$\begin{aligned}
|S| &= m^n (n+1)^r 2^{-n(\log m - H^*)} \\
&= (n+1)^r 2^{-nH^*}
\end{aligned}$$

and therefore, in general, $|S| \leq (n+1)^r 2^{-nH^*)}$.

■

Problem 0.4. *Differential entropy of the multivariate Gaussian*

$$\phi(x) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp -\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)$$

Solution.

$$\begin{aligned}
h(x) &= -\int \phi(x) \log \phi(x) dx \\
&= \int \phi(x) \left[\frac{n}{2}\log(2\pi e) + \frac{1}{2}\log|\Sigma| + \frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right] dx \\
&= \frac{n}{2}\log(2\pi e) + \frac{1}{2}\log|\Sigma| + \mathbf{E}\left[\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right] \\
&= \frac{n}{2}\log(2\pi e) + \frac{1}{2}\log|\Sigma|
\end{aligned}$$

■