Model Answers to Exercise 12 of May 18, 2017

http://www.isi.ee.ethz.ch/teaching/courses/it2.html

Problem 1 — *Large Deviations*

By Sanov's theorem, we have

$$\lim_{n\to\infty} \frac{1}{n} \log \Pr\left[\frac{1}{n}\sum_{k=1}^{n} X_k \geq \alpha\right] = -\inf_{\tilde{Q}\in\mathcal{F}} D(\tilde{Q}\|Q),$$

where Q denotes the geometric distribution and $\mathcal{F}$ is defined as

$$\mathcal{F} \triangleq \left\{\tilde{Q} \in \mathcal{P}(\mathbb{N}) : \sum_{i=1}^{\infty} i\tilde{Q}(i) \geq \alpha\right\}.$$

Since $\mathsf{E}_Q[X] = \frac{1}{p} < \alpha$, $Q \notin \mathcal{F}$ and we know from class that $\inf_{\tilde{Q}\in\mathcal{F}} D(\tilde{Q}\|Q) = D(Q^*\|Q)$ with

$$Q^*(i) = c \cdot Q(i)e^{-\lambda i} = c \cdot (1-p)^{i-1}pe^{-\lambda i} = c(1-p)^{-1}p \cdot \left((1-p)e^{-\lambda}\right)^i,$$

where λ is chosen such that $\mathsf{E}_{Q^*}[X] = \alpha$. Note that Q^* is again a geometric distribution, i.e.,

$$Q^*(i) = (1-q)^{i-1}q$$

with $q = 1-(1-p)e^{-\lambda}$. Comparing $\mathsf{E}_{Q^*}[X] = \alpha$ and $\mathsf{E}_{Q^*}[X] = \frac{1}{q}$, we see that $q = \frac{1}{\alpha}$. It remains to compute $D(Q^*\|Q)$:

$$\begin{aligned}
D(Q^*\|Q) &= \sum_{i=1}^{\infty} Q^*(i) \log \frac{Q^*(i)}{Q(i)} \\
&= \sum_{i=1}^{\infty} (1-q)^{i-1}q \log \frac{(1-q)^{i-1}q}{(1-p)^{i-1}p} \\
&= \sum_{i=1}^{\infty} (1-q)^{i-1}q \log\left(\frac{(1-p)q}{p(1-q)} \cdot \left(\frac{1-q}{1-p}\right)^i\right) \\
&= \left(\log \frac{(1-p)q}{p(1-q)}\right) \cdot \sum_{i=1}^{\infty}(1-q)^{i-1}q + \left(\log\frac{1-q}{1-p}\right) \cdot \sum_{i=1}^{\infty} i(1-q)^{i-1}q \\
&= \log\frac{(1-p)q}{p(1-q)} + \frac{1}{q}\log\frac{1-q}{1-p}.
\end{aligned}$$

Using $q = \frac{1}{\alpha}$, we obtain

$$\lim_{n\to\infty} \frac{1}{n} \log \Pr\left[\frac{1}{n}\sum_{k=1}^{n} X_k \geq \alpha\right] = -D(Q^*\|Q) = -\log\frac{(\alpha-1)^{\alpha-1}}{\alpha^{\alpha}(1-p)^{\alpha-1}p}.$$

Problem 2 ***Sanov-Type Theorem for the Size of Type Classes***

A sequence $\mathbf{x} \in \mathcal{X}^n$ satisfies $\frac{1}{n}\sum_{k=1}^{n} g(x_k) \geq \alpha$ if and only if its type satisfies $\sum_{x\in\mathcal{X}} P_{\mathbf{x}}(x)\, g(x) \geq \alpha$, so the size of the set $\bigl\{\mathbf{x} \in \mathcal{X}^n : \frac{1}{n}\sum_{k=1}^{n} g(x_k) \geq \alpha\bigr\}$ is equal to $\bigl|\mathcal{T}^n(\mathcal{P}_n(\mathcal{X}) \cap \mathcal{F})\bigr| = \bigl|\mathcal{T}^n(\mathcal{F})\bigr|$ with

$$\mathcal{F} = \left\{ Q \in \mathcal{P}(\mathcal{X})\colon \sum_{x\in\mathcal{X}} Q(x)\, g(x) \geq \alpha \right\},$$

and the claim about H^* indeed follows from the theorem.

We first prove the upper bound of the theorem. If $\mathcal{F} \cap \mathcal{P}_n(\mathcal{X})$ is empty, then the upper bound is trivially true. Otherwise,

$$\begin{aligned}
\bigl|\mathcal{T}^n(\mathcal{F})\bigr| &= \bigl|\mathcal{T}^n\bigl(\mathcal{F} \cap \mathcal{P}_n(\mathcal{X})\bigr)\bigr| \\
&= \sum_{P \in \mathcal{F}\cap\mathcal{P}_n(\mathcal{X})} \bigl|\mathcal{T}^n(P)\bigr| \\
&\overset{\text{(i)}}{\leq} \sum_{P \in \mathcal{F}\cap\mathcal{P}_n(\mathcal{X})} \max_{\tilde{P} \in \mathcal{F}\cap\mathcal{P}_n(\mathcal{X})} \bigl|\mathcal{T}^n(\tilde{P})\bigr| \\
&= \bigl|\mathcal{F} \cap \mathcal{P}_n(\mathcal{X})\bigr| \cdot \max_{\tilde{P} \in \mathcal{F}\cap\mathcal{P}_n(\mathcal{X})} \bigl|\mathcal{T}^n(\tilde{P})\bigr| \\
&\leq \bigl|\mathcal{P}_n(\mathcal{X})\bigr| \cdot \max_{\tilde{P} \in \mathcal{F}\cap\mathcal{P}_n(\mathcal{X})} \bigl|\mathcal{T}^n(\tilde{P})\bigr| \\
&\leq (n+1)^{|\mathcal{X}|} \max_{\tilde{P} \in \mathcal{F}\cap\mathcal{P}_n(\mathcal{X})} \bigl|\mathcal{T}^n(\tilde{P})\bigr| \\
&\leq (n+1)^{|\mathcal{X}|} \max_{\tilde{P} \in \mathcal{F}\cap\mathcal{P}_n(\mathcal{X})} e^{n\,\mathsf{H}(\tilde{P})} \\
&\leq (n+1)^{|\mathcal{X}|} \sup_{Q\in\mathcal{F}} e^{n\,\mathsf{H}(Q)} \\
&= (n+1)^{|\mathcal{X}|}\, e^{n \sup_{Q\in\mathcal{F}} \mathsf{H}(Q)},
\end{aligned}$$

where (i) is well-defined since the set $\mathcal{F} \cap \mathcal{P}_n(\mathcal{X})$ is finite and nonempty.

For the lower bound, assume that there exists a sequence $\{P_n \in \mathcal{F} \cap \mathcal{P}_n(\mathcal{X})\}$ satisfying

$$\liminf_{n\to\infty} \mathsf{H}(P_n) = \sup_{Q\in\mathcal{F}} \mathsf{H}(Q).$$

Then,

$$\begin{aligned}
\bigl|\mathcal{T}^n(\mathcal{F})\bigr| &= \bigl|\mathcal{T}^n\bigl(\mathcal{F} \cap \mathcal{P}_n(\mathcal{X})\bigr)\bigr| \\
&= \sum_{P \in \mathcal{F}\cap\mathcal{P}_n(\mathcal{X})} \bigl|\mathcal{T}^n(P)\bigr| \\
&\overset{\text{(i)}}{\geq} \bigl|\mathcal{T}^n(P_n)\bigr| \\
&\geq \frac{1}{(n+1)^{|\mathcal{X}|}}\, e^{n\,\mathsf{H}(P_n)},
\end{aligned}$$

where (i) holds because we drop all elements of the sum apart from one. Taking the logarithm, dividing by n, and taking the limit inferior yields

$$\liminf_{n\to\infty} \frac{1}{n} \log\bigl|\mathcal{T}^n(\mathcal{F})\bigr| \geq \liminf_{n\to\infty} \left\{ \frac{-|\mathcal{X}| \log(n+1)}{n} + \mathsf{H}(P_n) \right\} = \liminf_{n\to\infty} \mathsf{H}(P_n) = \sup_{Q\in\mathcal{F}} \mathsf{H}(Q).$$

The final statement now follows because the upper bound implies

$$\limsup_{n\to\infty} \frac{1}{n} \log\bigl|\mathcal{T}^n(\mathcal{F})\bigr| \leq \sup_{Q\in\mathcal{F}} \mathsf{H}(Q).$$

a) Since $D(\tilde{Q}\|Q_{\text{pass}})$ is infinite if there exists a symbol $z \in \mathbb{N}_0$ with $\tilde{Q}(z) > 0$ and $Q_{\text{pass}}(z) = 0$, we only consider PMFs $\tilde{Q}$ with $\operatorname{supp}(\tilde{Q}) \subseteq \{0, 10\}$. Let $\tilde{Q}(10) = 1 - \tilde{Q}(0) = p$. Then,

$$\mathsf{E}_{\tilde{Q}}[Z] = 10p + 0 \cdot (1-p) = 10p \overset{!}{\geq} 4,$$

which implies that $p \geq \frac{2}{5}$. Next, we show that $p^* = \frac{2}{5}$ minimizes $D(\tilde{Q}\|Q_{\text{pass}})$. Let Q^* be the PMF corresponding to p^*, let p be any value greater than or equal to $\frac{2}{5}$, and let $\tilde{Q}$ be the PMF corresponding to p. Then, $Q^* = \frac{1}{5p-1}\tilde{Q} + \frac{5p-2}{5p-1}Q_{\text{pass}}$ holds, and because $D(\tilde{Q}\|Q_{\text{pass}})$ is convex in $\tilde{Q}$, we have

$$\begin{aligned} D(Q^*\|Q_{\text{pass}}) &= D\left(\frac{1}{5p-1}\tilde{Q} + \frac{5p-2}{5p-1}Q_{\text{pass}}\middle\| Q_{\text{pass}}\right) \\ &\leq \frac{1}{5p-1}D(\tilde{Q}\|Q_{\text{pass}}) + \frac{5p-2}{5p-1}D(Q_{\text{pass}}\|Q_{\text{pass}}) \\ &= \frac{1}{5p-1}D(\tilde{Q}\|Q_{\text{pass}}) \\ &\leq D(\tilde{Q}\|Q_{\text{pass}}), \end{aligned}$$

which implies that $p^* = \frac{2}{5}$ is indeed optimal. Consequently,

$$\inf_{\tilde{Q}\in\mathcal{P}(\mathbb{N}_0)\colon\ \mathsf{E}_{\tilde{Q}}[Z]\geq 4} D(\tilde{Q}\|Q_{\text{pass}}) = D(Q^*\|Q_{\text{pass}}) = \frac{3}{5}\log\frac{\frac{3}{5}}{\frac{4}{5}} + \frac{2}{5}\log\frac{\frac{2}{5}}{\frac{1}{5}} = \frac{1}{5}\log\frac{27}{16} \approx 0.105 \text{ nats}.$$

b) Again note that $D(\tilde{Q}\|Q_{\text{run}})$ is infinite if there is any symbol $z \in \mathbb{N}_0$ with $\tilde{Q}(z) > 0$ and $Q_{\text{run}}(z) = 0$. Hence, only consider $z \in \{1, 2, 3, 4\}$. But a distribution on $\{1, 2, 3, 4\}$ can only have an expectation greater than or equal to 4 if the symbols 1, 2, and 3 have zero probability! Thus,

$$Q^*(z) = \begin{cases} 1 & \text{for } z = 4, \\ 0 & \text{otherwise.} \end{cases}$$

The relative entropy is then

$$D(Q^*\|Q_{\text{run}}) = 1 \cdot \log\frac{1}{\frac{1}{4}} = \log 4 \approx 1.39 \text{ nats}.$$

c) Sanov's theorem tells us that the probability of this event is about

$$\exp\left(-n \inf_{\tilde{Q}\in\mathcal{P}(\mathbb{N}_0)\colon\ \mathsf{E}_{\tilde{Q}}[Z]\geq 4} D(\tilde{Q}\|Q_{\text{pass}})\right) = \left(\frac{16}{27}\right)^{n/5} \approx 0.90^n.$$

d) Similarly, Sanov's theorem tells us that the probability of this event is about

$$\exp\left(-n \inf_{\tilde{Q}\in\mathcal{P}(\mathbb{N}_0)\colon\ \mathsf{E}_{\tilde{Q}}[Z]\geq 4} D(\tilde{Q}\|Q_{\text{run}})\right) = \left(\frac{1}{4}\right)^{n} = 0.25^n.$$

So clearly the strategy of only passing has a much higher probability of success.

Problem 4 ***Hypothesis Testing***

Applying Stein's lemma, we obtain that the optimal error exponent for $\Pr(\text{Decide } H_1|H_0 \text{ true})$ subject to $\Pr(\text{Decide } H_0|H_1 \text{ true}) \leq \frac{1}{2}$ is given by

$$\begin{aligned} D(Q\|P) &= \frac{1}{4}\log\frac{\frac{1}{4}}{\frac{1}{2}} + \frac{1}{4}\log\frac{\frac{1}{4}}{\frac{1}{4}} + \frac{1}{2}\log\frac{\frac{1}{2}}{\frac{1}{4}} \\ &= -\frac{1}{4} + 0 + \frac{1}{2} \\ &= \frac{1}{4} \text{ bits.} \end{aligned}$$

