6.842 Randomness and Computation April 3, 2022

Homework 2

Yuchong Pan *MIT ID: 911346847*

1. *Collaborators and sources:* Guanghao Ye.

2. (a) *Collaborators and sources:* Guanghao Ye.

Proof. Let $\{x, y\} \subset A$ be such that $x \neq y$. Then for any pairwise independent hash function $h \in B$,

$$(h(x), h(y)) \in_U T^2.$$

Therefore,

$$\mathop{\mathbb{P}}_{h \in_U B}[h(x) = h(y)] = \sum_{z \in T} \mathop{\mathbb{P}}_{h \in_U B}[(h(x), h(y)) = (z, z)] = |T| \cdot \frac{1}{|T^2|} = t \cdot \frac{1}{t^2} = \frac{1}{t}. \tag{1}$$

It follows that

$$\begin{aligned}
\mathop{\mathbb{E}}_{h \in_U B}[\#\text{ colliding pairs for } h] &= \mathop{\mathbb{E}}_{h \in_U B}\left[\sum_{\substack{\{x,y\} \subset A \\ x \neq y}} \mathbb{1}_{\{x,y\} \text{ is a colliding pair for } h}\right] \\
&= \sum_{\substack{\{x,y\} \subset A \\ x \neq y}} \mathop{\mathbb{E}}_{h \in_U B}\left[\mathbb{1}_{\{x,y\} \text{ is a colliding pair for } h}\right] \\
&= \sum_{\substack{\{x,y\} \subset A \\ x \neq y}} \mathop{\mathbb{P}}_{h \in_U B}[\{x, y\} \text{ is a colliding pair for } h] \\
&= \sum_{\substack{\{x,y\} \subset A \\ x \neq y}} \mathop{\mathbb{P}}_{h \in_U B}[h(x) = h(y)] \\
&= |\{\{x, y\} \subset A : x \neq y\}| \cdot \frac{1}{t} \\
&= \binom{|A|}{2} \cdot \frac{1}{t} \\
&= \binom{n}{2} \cdot \frac{1}{t}.
\end{aligned}$$

This completes the proof. $\square$

(b) *Collaborators and sources:* Guanghao Ye.

Proof. Let $p = (p_i)_{i \in A}$ be a distribution over A such that $c(p) \le (1+\varepsilon^2)/|A|$ for some $\varepsilon > 0$. Then $\sum_{i \in A} p_i = 1$ and $\sum_{i \in A} p_i^2 \le (1+\varepsilon^2)/|A|$. Therefore,

$$\begin{aligned}
\|p - U_A\|_1 &\le \sqrt{|A|}\,\|p - U_A\|_2 && \text{(Cauchy-Schwarz inequality)} \\
&= \sqrt{|A|}\sqrt{\sum_{i \in A}\left(p_i - \frac{1}{|A|}\right)^2} \\
&= \sqrt{|A|}\sqrt{\sum_{i \in A}\left(p_i^2 - \frac{2p_i}{|A|} + \frac{1}{|A|^2}\right)} \\
&= \sqrt{|A|}\sqrt{\sum_{i \in A} p_i^2 - \frac{2}{|A|}\sum_{i \in A} p_i + \sum_{i \in A}\frac{1}{|A|^2}} \\
&\le \sqrt{|A|}\sqrt{\frac{1+\varepsilon^2}{|A|} - \frac{2}{|A|}\cdot 1 + |A| \cdot \frac{1}{|A|^2}} \\
&= \sqrt{|A|}\sqrt{\frac{1+\varepsilon^2}{|A|} - \frac{2}{|A|} + \frac{1}{|A|}} \\
&= \sqrt{|A| \cdot \frac{1+\varepsilon^2 - 2 + 1}{|A|}} \\
&= \sqrt{\varepsilon^2} \\
&= \varepsilon.
\end{aligned}$$

This completes the proof. □

(c) *Collaborators and sources:* Guanghao Ye.

Proof. Let q be a distribution over $B \times T$ be defined as in the problem. Let $x, y \in A$. If $x = y$, then $h(x) = h(y)$ for any $h \in B$. If $x \neq y$, then (1) implies that for any $h \in B$,

$$\mathop{\mathbb{P}}_{x,y \in_U W}[h(x) = h(y) \mid x \neq y] = \frac{1}{t} = \frac{1}{|T|}.$$

For any set Ω,

$$\begin{aligned}
\mathop{\mathbb{P}}_{\omega_1,\omega_2 \in_U \Omega}[\omega_1 = \omega_2] &= \sum_{\omega \in \Omega} \mathop{\mathbb{P}}_{\omega_1,\omega_2 \in_U \Omega}[\omega_1 = \omega_2 = \omega] \\
&= \sum_{\omega \in \Omega} \mathop{\mathbb{P}}_{\omega_1 \in_U \Omega}[\omega_1 = \omega] \mathop{\mathbb{P}}_{\omega_2 \in_U \Omega}[\omega_2 = \omega] && \text{(independence)} \\
&= |\Omega| \cdot \frac{1}{|\Omega|} \cdot \frac{1}{|\Omega|} \\
&= \frac{1}{|\Omega|}.
\end{aligned}$$

This implies that $\mathbb{P}_{h_1,h_2 \in_U B}[h_1 = h_2] = 1/|B|$ and that $\mathbb{P}_{x_1,x_2 \in_U W}[x_1 = x_2] = 1/|W|$. Fix $h \in B$. Then

$$\begin{aligned}
\mathop{\mathbb{P}}_{x_1,x_2 \in_U W}[h(x_1) = h(x_2)] &= \mathop{\mathbb{P}}_{x_1,x_2 \in_U W}[x_1 = x_2] \mathop{\mathbb{P}}_{x_1,x_2 \in_U W}[h(x_1) = h(x_2) \mid x_1 = x_2] + \\
&\qquad \mathop{\mathbb{P}}_{x_1,x_2 \in_U W}[x_1 \neq x_2] \mathop{\mathbb{P}}_{x_1,x_2 \in_U W}[h(x_1) = h(x_2) \mid x_1 \neq x_2] \\
&\leq \frac{1}{|W|} \cdot 1 + 1 \cdot \frac{1}{|T|} \\
&= \frac{1}{|W|} + \frac{1}{|T|}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
c(q) &= \mathop{\mathbb{P}}_{\langle h_1,y_1\rangle,\langle h_2,y_2\rangle \in_q B \times T}[\langle h_1, y_1\rangle = \langle h_2, y_2\rangle] \\
&= \mathop{\mathbb{P}}_{\substack{h_1,h_2 \in_U B \\ x_1,x_2 \in_U W}}[h_1 = h_2, h_1(x_1) = h_2(x_2)] \\
&= \mathop{\mathbb{P}}_{h_1,h_2 \in_U B}[h_1 = h_2] \mathop{\mathbb{P}}_{\substack{h_1,h_2 \in_U B \\ x_1,x_2 \in_U W}}[h_1(x_1) = h_2(x_2) \mid h_1 = h_2] && \text{(independence)} \\
&= \frac{1}{|B|} \mathop{\mathbb{P}}_{\substack{h \in B \\ x_1,x_2 \in_U W}}[h(x_1) = h(x_2) \mid h] \\
&\leq \frac{1}{|B|}\left(\frac{1}{|W|} + \frac{1}{|T|}\right) \\
&= \frac{1}{|B|} \cdot \frac{|T|/|W| + 1}{|T|} \\
&= \frac{1 + |T|/|W|}{|B| \cdot |T|} \\
&= \frac{1 + |T|/|W|}{|B \times T|}.
\end{aligned}$$

This completes the proof. ∎

(d) *Collaborators and sources:* Guanghao Ye.

Proof. Note that it follows from the same argument of part (b) that for any distribution μ over any finite set Ω, if $c(\mu) \leq (1+\varepsilon^2)/|\Omega|$ for some $\varepsilon > 0$, then $\|\mu - U_\Omega\|_1 \leq \varepsilon$. Let $\Omega = B \times T$. Let $\varepsilon = \sqrt{|T|/|W|} > 0$. Then $|T|/|W| = \varepsilon^2$. By part (c),

$$c(q) \leq \frac{1 + |T|/|W|}{|B \times T|} = \frac{1+\varepsilon^2}{|\Omega|}.$$

Since q is a distribution over $B \times T = \Omega$, then

$$\|q - U_{B\times T}\|_1 = \|q - U_\Omega\|_1 \leq \varepsilon = \sqrt{|T|/|W|}.$$

This completes the proof. $\square$