

MCMT Homework 9

Shun Zhang

Exercise 9.1

Let Y_t denote the number of coordinates that have been selected after t rounds. Define the distinguishing statistic $f : \{0, 1\}^n \rightarrow R$ by $f(x) = \sum_{i=1}^n x_i$.

1. The probability that one coordinate is not selected in t steps is $p = (1 - \frac{1}{n})^t$. The expectation of the number of coordinates that are not selected in t steps is $E(n - Y_t) = np = n(1 - \frac{1}{n})^t$. So $E(Y_t) = n - n(1 - \frac{1}{n})^t$.

2. Let $I_i(t)$ be the indicator that i -th coordinate has been selected at t -th step.

$$E(I_i(t)I_j(t)) = 2(1 - (1 - \frac{1}{n})^t) - (1 - (1 - \frac{2}{n})^t) = 1 - 2(1 - \frac{1}{n})^t + (1 - \frac{2}{n})^t.$$

$$E(I_i(t))E(I_j(t)) = (1 - (1 - \frac{1}{n})^t)^2.$$

$$\text{Cov}(I_i(t), I_j(t)) = E(I_i(t)I_j(t)) - E(I_i(t))E(I_j(t)) = (1 - \frac{2}{n})^t - (1 - \frac{1}{n})^{2t} \leq 0.$$

$$\text{Var}(Y_t) = \text{Var}(\sum_i I_i(t)) = \sum_i (\text{Var}(I_i(t))) + \sum_{i \neq j} \text{Cov}(I_i(t), I_j(t)) \leq \sum_i (\text{Var}(I_i(t))) \leq np(1 - p) \leq \frac{n}{4}.$$

3. $E_0(X_{ti}|Y_t) = \frac{Y_t}{n} \frac{1}{2}$. This is the probability that x_i is chosen, and times the probability that it is set to be 1.

$$E_0(f(X_t)|Y_t) = \sum_i E_0(X_{ti}|Y_t) = \frac{Y_t}{2}.$$

$$E_0(f(X_t)) = \sum_y E_0(f(X_t)|Y_t = y)P(Y_t = y) = \sum_y \frac{y}{2}P(Y_t = y) = \frac{E(Y_t)}{2}.$$

4. $\text{Var}_0(f(X_t)|Y_t) = Y_t \frac{1}{2} (1 - \frac{1}{2}) = \frac{1}{4} Y_t$, as $f(X_t)$ given Y_t is a Binomial distribution. $E_0 \text{Var}_0(f(X_t)|Y_t) = \frac{1}{4} E_0(Y_t)$.

$$E_0(f(X_t)|Y_t) = \frac{Y_t}{2}. \quad \text{Var}_0 E_0(f(X_t)|Y_t) = \text{Var}_0 \frac{Y_t}{2} = \frac{1}{4} \text{Var}_0(Y_t).$$

$$\text{Var}_0(f(X_t)) = E_0 \text{Var}_0(f(X_t)|Y_t) + \text{Var}_0(E_0(f(X_t)|Y_t)) = \frac{1}{4} (E(Y_t) + \text{Var}(Y_t)).$$

By substituting by the results we have, $E(Y_t) + \text{Var}(Y_t) = n - n(1 - \frac{1}{n})^t + (n^2 - n) \text{cov} + n(1 - \frac{1}{n})^t (1 - (1 - \frac{1}{n})^t) = n + (n^2 - n) \text{cov} - n(1 - \frac{1}{n})^t < n$, where cov is $\text{Cov}(I_i(t), I_j(t))$ for $i \neq j$, which we have shown to be negative.

So, $\text{Var}_0(f(X_t)) < \frac{1}{4}n$.

5. f on π is a Binomial distribution. So $E_\pi f = \frac{n}{2}$, $\text{Var}_\pi f = n \frac{1}{2} (1 - \frac{1}{2}) = \frac{n}{4}$.

6. Consider two distributions $P^t(0, \cdot)$ and π .

$$\sigma^2 \leq (\frac{n}{4} + \frac{n}{4})/2 = \frac{n}{4}.$$

$$\Delta = |\frac{EY_t}{2} - \frac{n}{2}| = \frac{n - EY_t}{2} = \frac{n(1 - \frac{1}{n})^t}{2}.$$

$$\text{By Lemma 9.7, } \|P^t(0, \cdot) - \pi\|_{TV} \geq \frac{\Delta^2}{4\sigma^2 + \Delta^2} = 1 - \frac{4\sigma^2}{4\sigma^2 + \Delta^2} \geq 1 - \frac{n}{n + \frac{n^2(1 - \frac{1}{n})^{2t}}{4}} = 1 - 8 \frac{1}{8 + 2n(1 - \frac{1}{n})^{2t}} = 1 - 8 \exp\{-\log(8 + 2n(1 - \frac{1}{n})^{2t})\}$$

We know $d(t) \geq \|P^t(0, \cdot) - \pi\|_{TV}$. Let $t = \frac{1}{2}n \log n - cn$.

$$\begin{aligned} d(\frac{1}{2}n \log n - cn) &\geq 1 - 8 \exp\{-\log(8 + 2n(1 - \frac{1}{n})^{n \log n - 2cn})\} \\ &\geq 1 - 8 \exp\{-2c + 1\} \end{aligned}$$