CS 541

HW 1

Chenxu Wang

10457625

Problem 2.

1. Let $y_1\ y_2 \cdots y_n$ be independant random variables.

2.

Since $Pr(y_i = 1) = 0.6$

$Pr(y_i = -1) = 0.4$

$X = y_1\ y_2 \cdots y_n$

$$E(X) = \sum_{i=1}^{n} y_i \cdot p_i = n \cdot 0.2$$

$$Var(X) = \sum_{i=1}^{n} p_i (y_i - \mu) = 0.96n$$

when the majority vote $X > 0$, we can get the label which is a digit.

$$|X - E(X)| \geq a$$

$$\Rightarrow E(X) - a \leq X \leq E(X) + a$$

Let $E(X) - a = 0$

$\therefore\ a = E(X) = 0.2n$

Using Chebyshev's inequality:

$$P\{|X - E(X)| \leq E(X)\} \geqslant \frac{Var(X)}{E^2(X)}$$

Since the confidence is as high as 0.99.

$$\therefore \quad \frac{Var(X)}{E^2(X)} = 1 - 0.99$$

$$\frac{0.16n}{0.04n^2} = 0.01$$

$$\Rightarrow \quad n = 2400$$

$\therefore$ When $n > 2400$, we can get right label with a confidence over 0.99.

3. Since $Z \geq 0$ isn't good condition when random variable in $\{0, 1\}$ for the first formula, therefore, we choose:

$$Pr[Z \leq (1-\alpha)\eta \cdot n] \leq e^{-\frac{\alpha^2 \eta n}{2}}$$

Since confidence as high as 0.99

$$e^{\frac{-\alpha^2 \eta n}{2}} = 1 - 0.99$$

since we need majority vote $Z > 0$

Let $(1-\alpha)\cdot\eta\cdot n = 0$

$$\therefore \quad \alpha = 1$$

So we have:

$$e^{\frac{-1^2 \cdot 0.6 \cdot n}{2}} = 0.01$$

$$n \gtrsim 15.3$$

Finally, we get $n = 16$.

4 ① Chernoff bound requires that the variates be independent. - a condition that Chebyshev's inequality doesn't require, although Chebyshev's inequality requires the variates to be pairwise independent.

② Given 1. 2. we find that chernoff gives a much stronger bound than Chebyshev. This might be because Chebyshev only uses pairwise independence whereas chernoff uses full independence.