Note that we assume $0 < \varepsilon < \frac{1}{3}$ in our paper.

# 1 Proof of Theorem 3

According to (3.1),(3.5),(3.6) in [1] (Chernoff inequality), we have:

$$Pr(\hat{p}_n - p > \varepsilon) \leq f^n(p, \varepsilon) \qquad (p < 1 - \varepsilon) \tag{1}$$

$$Pr(p - \hat{p}_n > \varepsilon) \leq f^n(1 - p, \varepsilon) \qquad (p > \varepsilon) \tag{2}$$

Note that:

$$Pr(|\hat{p}_n - p| > \varepsilon) = Pr(\hat{p}_n - p > \varepsilon) + Pr(p - \hat{p}_n > \varepsilon) \tag{3}$$

therefore,

$$Pr(|\hat{p}_n - p| > \varepsilon) \leq f^n(p, \varepsilon) + f^n(1 - p, \varepsilon) \qquad (\varepsilon < p < 1 - \varepsilon) \tag{4}$$

When $p \leq \varepsilon$,

$$Pr(p - \hat{p}_n > \varepsilon) = 0 \tag{5}$$

therefore,

$$Pr(|\hat{p}_n - p| > \varepsilon) = Pr(\hat{p}_n - p > \varepsilon) \leq f^n(p, \varepsilon) \qquad (p < 1 - \varepsilon \wedge p \leq \varepsilon) \tag{6}$$

Similarly, when $p \geq 1 - \varepsilon$,

$$Pr(|\hat{p}_n - p| > \varepsilon) = Pr(p - \hat{p}_n > \varepsilon) \leq f^n(1 - p, \varepsilon) \qquad (p > \varepsilon \wedge p \geq 1 - \varepsilon) \tag{7}$$

When $\frac{1-\varepsilon}{2} \leq p \leq \frac{1+\varepsilon}{2}$, we directly use the Okamoto bound.

# 2 Proof of Theorem 4

**Lemma 1:** $f(p, \varepsilon)$ is monotonically increasing when $p \in [0, \frac{1-\varepsilon}{2}]$, and monotonically decreasing when $p \in [\frac{1}{2}, 1 - \varepsilon)$.

*Proof.* As $f(0, \varepsilon) = 0$, and when $0 < p < 1 - \varepsilon$, $f(p, \varepsilon) > 0$, therefore, we only need to consider $g(p, \varepsilon) \triangleq \ln(f(p, \varepsilon)) = (p + \varepsilon)(\ln(p) - \ln(p + \varepsilon)) + (1 - p - \varepsilon)(\ln(1 - p) - \ln(1 - p - \varepsilon))$ when $p \in (0, 1 - \varepsilon)$.

$$g'_p(p, \varepsilon) = \ln(p) - \ln(p + \varepsilon) + \frac{\varepsilon}{p} - \ln(1 - p) + \ln(1 - p - \varepsilon) + \frac{\varepsilon}{1 - p} \tag{8}$$

$$g''_p(p, \varepsilon) = -\varepsilon^2\left(\frac{1}{p^2(p + \varepsilon)} + \frac{1}{(1 - p)^2(1 - p - \varepsilon)}\right) < 0 \tag{9}$$

Therefore, $g'_p(p, \varepsilon)$ is monotonically decreasing when $p \in (0, 1 - \varepsilon)$.

Moreover,

$$g'_p\left(\frac{1 - \varepsilon}{2}, \varepsilon\right) = 2 \cdot \ln\left(\frac{1 - \varepsilon}{1 + \varepsilon}\right) + \frac{4\varepsilon}{1 - \varepsilon^2} > 0 \tag{10}$$

$$g_p'(\frac{1}{2},\varepsilon) = \ln(\frac{1-2\varepsilon}{1+2\varepsilon}) + 4\varepsilon < 0 \tag{11}$$

Therefore, when $p \in (0, \frac{1-\varepsilon}{2}]$, $g_p'(p,\varepsilon) > 0$, and when $p \in [\frac{1}{2}, 1-\varepsilon)$, $g_p'(p,\varepsilon) < 0$. ☐

**Lemma 2:** When $0 < \varepsilon < 1$ and $p < 1-\varepsilon$, $f(p,\varepsilon) < C(\varepsilon) < 1$.

*Proof.* It is trivial that $C(\varepsilon) < 1$. To prove $f(p,\varepsilon) < C(\varepsilon)$, we take logarithm on both sides, we only need to prove:

$$h(p,\varepsilon) \triangleq (p+\varepsilon)\ln(\frac{p+\varepsilon}{p}) + (1-p-\varepsilon)\ln(\frac{1-p-\varepsilon}{1-p}) - 2\varepsilon^2 > 0 \tag{12}$$

It follows immediately from the fact that $h(p,0) = 0$ and

$$h_\varepsilon'(p,\varepsilon) = \ln(1 + \frac{\varepsilon}{-(p-\frac{1-\varepsilon}{2})^2 + \frac{(1-\varepsilon)^2}{4}}) - 4\varepsilon > \ln(1+\frac{4\varepsilon}{(1-\varepsilon)^2}) - 4\varepsilon > 0 \tag{13}$$

☐

**Theorem 4:**
(1) For any $\varepsilon$ and $n$, when $0 \le p_1 < p_2 \le \frac{1}{2}$, $F(p_1,\varepsilon,n) \le F(p_2,\varepsilon,n)$; when $\frac{1}{2} \le p_1 < p_2 \le 1$, $F(p_1,\varepsilon,n) \ge F(p_2,\varepsilon,n)$.
(2) For any $\varepsilon$ and $p$, when $n_1 > n_2$, $F(p,\varepsilon,n_1) < F(p,\varepsilon,n_2)$.

*Proof.* For (2), from Lemma 2, we know that $f(p,\varepsilon) < C(\varepsilon) < 1$, therefore, for any $\varepsilon$ and $p$, when $n_1 > n_2$, $F(p,\varepsilon,n_1) < F(p,\varepsilon,n_2)$.

For (1), note that $F(p,\varepsilon,n) = F(1-p,\varepsilon,n)$, therefore, by symmetry, we only need to prove $F(p_1,\varepsilon,n) \le F(p_2,\varepsilon,n)$ when $0 \le p_1 < p_2 \le \frac{1}{2}$. As $F(p,\varepsilon,n)$ is a piecewise continuous function (with regard to $p$) on $[0,\frac{1}{2}]$, we only need to prove it in three continuous intervals and two discontinuity points.

When $p \le \varepsilon$, we know $f(p,\varepsilon)$ is monotonically increasing when $p \in [0,\varepsilon]$ (Lemma 1). Therefore, $F(p,\varepsilon,n) = f^n(p,\varepsilon)$ is also monotonically increasing when $p \in [0,\varepsilon]$.

At the discontinuity point $p = \varepsilon$, $F(\varepsilon+,\varepsilon,n) = f^n(\varepsilon,\varepsilon) + (1-\varepsilon)^n > f^n(\varepsilon,\varepsilon) = F(\varepsilon,\varepsilon,n)$.

When $\varepsilon < p < \frac{1-\varepsilon}{2}$, as $f(p,\varepsilon)$ and $f(1-p,\varepsilon)$ are both monotonically increasing when $p \in (\varepsilon, \frac{1-\varepsilon}{2})$ (Lemma 1). Therefore, $F(p,\varepsilon,n) = f^n(p,\varepsilon) + f^n(1-p,\varepsilon)$ is also monotonically increasing when $p \in (\varepsilon, \frac{1-\varepsilon}{2})$.

At the discontinuity point $p = \frac{1-\varepsilon}{2}$, we know from Lemma 2 that $F(\frac{1-\varepsilon}{2}-,\varepsilon,n) = f^n(\frac{1-\varepsilon}{2},\varepsilon) + f^n(\frac{1+\varepsilon}{2},\varepsilon) < 2\cdot C^n(\varepsilon) = F(\frac{1-\varepsilon}{2},\varepsilon,n)$.

When $\frac{1-\varepsilon}{2} \le p \le \frac{1}{2}$, $F(p,\varepsilon,n) = 2\cdot C^n(\varepsilon)$ is independent of $p$. ☐

## 3 Proof of Theorem 5

When $ub \leq \varepsilon$, $F(p, \varepsilon, n) = f^n(p, \varepsilon)$. According to Theorem 4, for any $\varepsilon, n$ and $p \in [lb, ub]$, $F(p, \varepsilon, n) \leq F(ub, \varepsilon, n)$. Therefore, when $n_1 = \lceil \log_{f(ub,\varepsilon)} \delta \rceil$, we have $F(p, \varepsilon, n_1) \leq F(ub, \varepsilon, n_1) = f^{n_1}(ub, \varepsilon) \leq \delta$ for all $p \in [lb, ub]$. Moreover, for any $n_2 < n_1$, $F(ub, \varepsilon, n_2) = f^{n_2}(ub, \varepsilon) > \delta$. Therefore, $n_1 = \lceil \log_{f(ub,\varepsilon)} \delta \rceil$ is the smallest $n$ such that $F(p, \varepsilon, n) \leq \delta$ for all $p \in [lb, ub]$. The case when $b \geq 1 - \varepsilon$ is similar.

When $[lb, ub] \cap [\frac{1-\varepsilon}{2}, \frac{1+\varepsilon}{2}] \neq \emptyset$ (i.e. $(\frac{1-\varepsilon}{2} \leq ub \leq \frac{1+\varepsilon}{2}) \vee (\frac{1-\varepsilon}{2} \leq lb \leq \frac{1+\varepsilon}{2}) \vee (ub \geq \frac{1+\varepsilon}{2} \wedge lb \leq \frac{1-\varepsilon}{2})$), for any $\varepsilon, n$, the maximum of $F(p, \varepsilon, n)$ when $p \in [lb, ub]$ is $2 \cdot C^n(\varepsilon)$. Therefore, $n = \lceil \frac{1}{2\varepsilon^2} \ln(\frac{2}{\delta}) \rceil$ is the smallest $n$ such that $F(p, \varepsilon, n) \leq \delta$ for all $p \in [lb, ub]$.

When $\varepsilon < ub < \frac{1-\varepsilon}{2}$, for any $\varepsilon, n$ and $p \in [lb, ub]$, $F(p, \varepsilon, n) \leq F(ub, \varepsilon, n) = f^n(ub, \varepsilon) + f^n(1 - ub, \varepsilon)$. When we perform binary search, the initial lower bound is 0, and the initial upper bound is $\lceil \frac{1}{2\varepsilon^2} \ln(\frac{2}{\delta}) \rceil$. It is obvious that $F(ub, \varepsilon, 0) = 2 > \delta$, therefore, the initial lower bound is valid. When $n = \lceil \frac{1}{2\varepsilon^2} \ln(\frac{2}{\delta}) \rceil$, $F(ub, \varepsilon, n) = f^n(ub, \varepsilon) + f^n(1 - ub, \varepsilon) < 2 \cdot C^n(\varepsilon) = \delta$ (Lemma 2), therefore, the initial upper bound is also valid. As the initial lower and upper bound are both valid, the binary search can find the smallest $n$ such that $F(p, \varepsilon, n) \leq \delta$ for all $p \in [lb, ub]$ (note that $F(p, \varepsilon, n)$ is monotonically decreasing with regard to $n$, see Theorem 4). The case when $\frac{1+\varepsilon}{2} < lb < 1 - \varepsilon$ is similar.

## References

[1] Chernoff, H.: A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. The Annals of Mathematical Statistics **23**(4), 493–507 (December 1952)