Probability and Computing, 2nd Edition

Solutions to Chapter 4: Chernoff and Hoeffding Bounds

Hahndeul Kim

July 2025

4.1

Let the number of games that Alice wins be X, where $X \sim B(n, 0.6)$. Alice will lose the tournament with probability $\Pr(X \leq \frac{n-1}{2})$. Now, let δ s. t. $(1-\delta) \times \frac{3n}{5} = \frac{n-1}{2}$ to obtain the tightest bound.
$\Pr(X \leq \frac{n-1}{2}) = \Pr(X \leq (1-\delta)\mathbf{E}[X]) \leq \exp(-\frac{3n}{5} \cdot \delta^2 \cdot \frac{1}{2})$
$= \exp(-\frac{1}{10}(\frac{1}{12}n + \frac{5}{6} + \frac{25}{12n})) \leq \exp(-\frac{1}{8})$ (AM-GM inequality).

4.2

With Markov's inequality, $\Pr(X \geq n/4) \leq (n/6)/(n/4) = 2/3$.
With Chebyshev's inequality, $\Pr(X \geq n/4) \leq \Pr(|X - n/6| \geq n/12) \leq \frac{\mathbf{Var}[X]}{(n/12)^2}$
$= \frac{144}{n^2} \times (n \cdot \frac{1}{6} \cdot \frac{5}{6}) = 20/n$.
To use Chernoff bounds, let $\delta = 1/2$. Then $\Pr(X \geq n/4) = \Pr(X \geq (1+\delta)\mathbf{E}[X])$
$\leq \left(\frac{e^{0.5}}{1.5^{1.5}}\right)^{n/6} = \left(\frac{e}{1.5^3}\right)^{n/12}$.

4.3

(a) Let $X \sim B(n,p)$. Then $M_X(t) = \mathbf{E}[e^{tX}] = \sum\limits_{i=0}^{n} e^{it} \Pr(X = i)$
$= \sum\limits_{i=0}^{n} e^{it} \binom{n}{i} p^i (1-p)^{n-i} = \sum\limits_{i=0}^{n} \binom{n}{i} (pe^t)^i (1-p)^{n-i} = (pe^t + 1 - p)^n$.
(b) $M_{X+Y}(t) = \mathbf{E}[e^{t(X+Y)}] = \mathbf{E}[e^{tX}e^{tY}] = \mathbf{E}[e^{tX}]\mathbf{E}[e^{tY}] = (pe^t + 1 - p)^{m+n}$.
(c) Since moment generating function uniquely determines the distribution, $X + Y \sim B(m+n, p)$.

4.4

Let the total number of heads be X, where $X \sim B(100, \frac{1}{2})$. Then we find $\Pr(X \geq 55) \approx 0.1841$.
From Chernoff bound, we find that $\Pr(X \geq (1 + \frac{1}{10})50) \leq \exp(-\frac{50}{3} \cdot \frac{1}{10^2}) = \exp(-\frac{1}{6}) \approx 0.8465$.
For $Y \sim B(1000, \frac{1}{2})$, $\Pr(Y \geq 550) \approx 0.0009$.
From Chernoff bound, we find that $\Pr(Y \geq (1 + \frac{1}{10})500) \leq \exp(-\frac{500}{3} \cdot \frac{1}{10^2}) = \exp(-\frac{5}{3}) \approx 0.1889$.

4.5

Let $Y = NX$, so that we aim to satisfy $\Pr(|Y - Np| > N\epsilon p) \leq \delta$. Consider that $\Pr(Y > Np(1+\epsilon)) < \exp(-Np \cdot \frac{\epsilon^2}{3})$, and $\Pr(Y < Np(1-\epsilon)) < \exp(-Np \cdot \frac{\epsilon^2}{2})$. Thus, we aim to satisfy $\exp(-Np \cdot \frac{\epsilon^2}{3}) + \exp(-Np \cdot \frac{\epsilon^2}{2}) \leq 2\exp(-Np \cdot \frac{\epsilon^2}{3}) \leq \delta$.

$\therefore N \geq \frac{3}{p\epsilon^2} \ln \frac{2}{\delta}$. With $\epsilon = 0.1$, $\delta = 0.05$ and $0.2 \leq p \leq 0.8$, $N \geq 1500 \ln 40 \approx 5533$.

4.6

(a) Let $X \sim B(1000000, 0.02)$. Then $\Pr(X \geq 40000) \leq e^{-20000/3}$.
(b) Set X and Y as given and choose k, l such that $l \leq k - 10000$ so that bounding $\Pr((X > k) \cap (Y < l))$ suffices. As examples, we choose $k = 15300$ and $l = 4900$ here. Since $X \sim B(510000, 0.02)$, $Y \sim B(490000, 0.02)$ and $X \perp\!\!\!\perp Y$, $\Pr((X > k) \cap (Y < l)) = \Pr(X > k)\Pr(Y < l) \leq e^{-10200/12} \times e^{-9800/8} = e^{-2025}$.

4.7

Recall that $M_X(t) = \prod_{i=1}^{n} (p_i e^t + (1 - p_i)) = \prod_{i=1}^{n} (1 + p_i(e^t - 1)) \leq \prod_{i=1}^{n} e^{p_i(e^t - 1)}$
$= e^{\mu(e^t - 1)}$ holds when X is the sum of Poisson trials ($\Pr(X_i = 1) = p_i$).
Let $t = \ln(1 + \delta)$ and follow the derivation of Chernoff bounds.
$\Pr(X \geq (1+\delta)\mu_H) \leq \frac{\mathbf{E}[e^{tX}]}{e^{t(1+\delta)\mu_H}} \leq \frac{e^{\mu(e^t - 1)}}{e^{t(1+\delta)\mu_H}} \leq \left(\frac{e^{e^t - 1}}{e^{t(1+\delta)}}\right)^{\mu_H} = \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu_H}$.
Similarly, let $t = \ln(1 - \delta)$ and prove the latter inequality.
$\Pr(X \leq (1-\delta)\mu_L) \leq \frac{\mathbf{E}[e^{tX}]}{e^{t(1-\delta)\mu_L}} \leq \frac{e^{\mu(e^t - 1)}}{e^{t(1-\delta)\mu_L}} \leq \left(\frac{e^{e^t - 1}}{e^{t(1-\delta)}}\right)^{\mu_L} = \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu_L}$. ■

4.8

For any permutation π produced with the given approach, $\Pr(f = \pi) = \prod_{i=1}^{n} \frac{1}{k+1-i}$ holds. Since the number of possible permutations is $\frac{k!}{(k-n)!} = \frac{1}{\Pr(f=\pi)}$, the given approach produces a permutation chosen uniformly at random from all permutations.
Now, let X_j be the number of black box calls to determine $f(j)$. Then $X_j \sim Geom(\frac{k+1-j}{k})$ holds. Thus, $\mathbf{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbf{E}[X_i] = \sum_{i=1}^{n} \frac{k}{k+1-i}$.
When $k = n$, $\mathbf{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \frac{n}{i} = nH(n) \approx n \ln n$.
Similarly, when $k = 2n$, $\mathbf{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \frac{2n}{n+i} = 2n(H(2n) - H(n)) \approx 2n \ln 2$. In this case, $\frac{2n+1-j}{2n} \geq \frac{2n+1-n}{2n} \geq \frac{1}{2}$.
Now, to derive the desired Chernoff bound, we first compute the moment generating function of $X = \sum_{i=1}^{n} X_j$. Let $p_i = \frac{2n+1-i}{2n}$. Since X_i are independent,
$\mathbf{E}[e^{tX}] = \prod_{i=1}^{n} \mathbf{E}[e^{tX_i}] = \prod_{i=1}^{n} \left(\prod_{j=1}^{\infty} (e^{tj} p_i (1-p_i)^{j-1}\right) = \prod_{i=1}^{n} \left(\frac{p_i}{1-p_i} \prod_{j=1}^{\infty} (e^t(1-p_i))^j\right)$.
Suppose that we choose t s. t. $0 < t < \ln 2$ when deriving the Chernoff bound.

Then $\mathbf{E}[e^{tX}] = \prod_{i=1}^{n} \frac{p_i e^t}{1-e^t(1-p_i)}$. Since $t > 0$, $\frac{\partial}{\partial p_i}\left(\frac{p_i e^t}{1-e^t(1-p_i)}\right) = \frac{1-e^t}{(1-e^t(1-p_i))^2} < 0$.
This leads to $\mathbf{E}[e^{tX}] = \prod_{i=1}^{n} \frac{p_i e^t}{1-e^t(1-p_i)} \le \left(\frac{\frac{1}{2}e^t}{1-\frac{1}{2}e^t}\right)^n$.
Now derive the desired Chernoff bound with $\Pr(X \ge 4n) \le \frac{\mathbf{E}[e^{tX}]}{e^{4nt}} \le \left(\frac{1}{(2-e^t)e^{3t}}\right)^n$.
Since the function $(2-e^t)e^{3t}$ has its maximum at $t = \ln\frac{3}{2}$ and $0 < \ln\frac{3}{2} < \ln 2$, we choose $t = \ln\frac{3}{2}$ for the tightest possible bound.
The desired bound would be $\Pr(X \ge 4n) \le \left(\frac{1}{(2-e^t)e^{3t}}\right)^n \Big|_{t=\ln\frac{3}{2}} = (\frac{16}{27})^n$.

4.9

(a) By Chebyshev's inequality, $\Pr[|\sum_{i=1}^{t} X_i - \mathbf{E}[X]| \ge \epsilon\mathbf{E}[X]] \le \frac{\mathbf{Var}[X]}{t(\epsilon\mathbf{E}[X])^2} = \frac{r^2}{t\epsilon^2}$.
Thus, setting t to satisfy $\frac{r^2}{t\epsilon^2} \le \delta$ suffices. This leads to $t \ge \frac{r^2}{\epsilon^2\delta}$, which proves the claim.
(b) Set $\delta = 1 - 3/4 = 1/4$. Then we get $t \ge \frac{4r^2}{\epsilon^2}$, which proves the claim.
(c) Let Y_i be indicator variables that are 1 if $|X_i - \mathbf{E}[X]| \ge \epsilon\mathbf{E}[X]$. Then let the median of Y_is be m, and bound the probability $\Pr(|m - \mathbf{E}[X]| \ge \epsilon\mathbf{E}[X])$.
Note that $\mathbf{E}[\sum_{i=1}^{t} Y_i] \le t/4$ by definition, and $|m - \mathbf{E}[X]| \ge \epsilon\mathbf{E}[X]$ holds only if $\sum_{i=1}^{t} Y_i \ge t/2$. Then, $\Pr(|m - \mathbf{E}[X]| \ge \epsilon\mathbf{E}[X]) \le \Pr\left(\sum_{i=1}^{t} Y_i \ge t/2\right)$. Let $Y = \sum_{i=1}^{t} Y_i$. Then $\Pr(Y \ge t/2) = \Pr\left(Y \ge (1 + (\frac{t}{2\mathbf{E}[Y]} - 1))\mathbf{E}[Y]\right)$
$\le \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^{\mathbf{E}[Y]} = (\frac{2e}{t})^{t/2} \times e^{-\mathbf{E}[Y]}\mathbf{E}[Y]^{t/2}$.
Since $\frac{\partial}{\partial\mathbf{E}[Y]}\left((\frac{2e}{t})^{t/2}e^{-\mathbf{E}[Y]}\mathbf{E}[Y]^{t/2}\right) = (\frac{2e}{t})^{t/2}e^{-\mathbf{E}[Y]}\mathbf{E}[Y]^{t/2-1}(t/2 - \mathbf{E}[Y]) > 0$, substitute $t/4$ for $\mathbf{E}[Y]$ to derive our bound. Thus, $\Pr(Y \ge t/2) \le (\frac{e}{4})^{t/4}$.
Here we need t that satisfies $(\frac{e}{4})^{t/4} \le \delta$, which leads to $t \ge \frac{4}{\ln\frac{4}{e}}\ln\frac{1}{\delta}$. Therefore, together with 4.9.(b), we only need $O(\log(1/\delta))$ estimates constructed from $O(r^2\log(1/\delta)/\epsilon^2)$ samples.

4.10

Let $X = \sum_{i=1}^{1000000} X_i$ where X_i denotes the winnings of the ith game.
Then by the Chernoff bound, $\Pr(X \ge 10000) \le \frac{\mathbf{E}[e^{tX}]}{e^{10000t}} = \left(\frac{\mathbf{E}[e^{tX_i}]^{100}}{e^t}\right)^{10000}$
$= \left(\frac{(167/200)e^{-t} + (4/25)e^{2t} + (1/200)e^{99t}}{e^{0.01t}}\right)^{1000000}$. Using graph software, you can choose $t = 0.0006$ and derive $\Pr(X \ge 10000) \le 0.0001606$.

4.11

Since $\mathbf{E}[X_i] = 1$, $\mathbf{E}[X] = n$. Thus, we bound $\Pr(X \geq (1+\delta)n)$ as $\Pr(X \geq (1+\delta)n) \leq \frac{\mathbf{E}[e^{tX}]}{e^{t(1+\delta)n}}$ with $t > 0$. $\mathbf{E}[e^{tX}] = \prod_{i=1}^{n} \mathbf{E}[e^{tX_i}] = \left(\frac{1}{3}(1+e^t+e^{2t})\right)^n$ leads to $\Pr(X \geq (1+\delta)n) \leq \left(\frac{1+e^t+e^{2t}}{3e^{t(1+\delta)}}\right)^n$. Although $t = \frac{\delta+\sqrt{4-3\delta^2}}{1-\delta}$ minimizes $\frac{1+e^t+e^{2t}}{3e^{t(1+\delta)}}$, it is too complex to be used as a generalized bound. Thus, we put $t = \ln(1+\delta)$ for simplicity and derive $\Pr(X \geq (1+\delta)n) \leq \left(\frac{3+3\delta+\delta^2}{3(1+\delta)^{(1+\delta)}}\right)^n$.
The Chernoff bound for $\Pr(X \leq (1-\delta)n)$ can also be derived in a similar way.

4.12

(a) We can think of X_i as the number of tails between $i-1$th head and ith head. Now, let Y_i be indicator variables that are 1 if ith flip is head. Then let $Y = \sum_{i=1}^{(1+\delta)2n} Y_i$, and derive the Chernoff bound as $\Pr(X \geq (1+\delta)2n) = \Pr(Y \leq n)$. Since $\mathbf{E}[Y] = (1+\delta)n$, $\Pr(Y \leq n) = \Pr(Y \leq (1-\frac{\delta}{1+\delta})\mathbf{E}[Y]) \leq e^{-\frac{1}{2}\mathbf{E}[Y](\frac{\delta}{1+\delta})^2} = e^{-\frac{n\delta^2}{2(1+\delta)}}$.
(b) Here, the moment generating function for X can be derived as $\mathbf{E}[e^{tX}] = \left(\frac{e^t}{2-e^t}\right)^n$ for $0 < t < \ln 2$ (refer to the solution for exercise 4.8).
Thus, $\Pr(X \geq (1+\delta)2n) \leq \frac{(\frac{e^t}{2-e^t})^n}{e^{t(1+\delta)2n}} = \left(\frac{1}{e^{t(1+2\delta)}(2-e^t)}\right)^n$. Since $e^{t(1+2\delta)}(2-e^t)$ is maximized at $t = \ln(\frac{1+2\delta}{1+\delta}) < \ln 2$, we choose it to derive the tightest bound. Therefore, $\Pr(X \geq (1+\delta)2n) \leq \left((\frac{1+\delta}{1+2\delta})^{1+2\delta}(1+\delta)\right)^n$.
(c) To compare two bounds, we inspect the sign of $e^{-\frac{\delta^2}{2(1+\delta)}} - (\frac{1+\delta}{1+2\delta})^{1+2\delta}(1+\delta)$. The simpler equivalent would be $(1+2\delta)\ln(1+2\delta) - (2+2\delta)\ln(1+\delta) - \frac{\delta^2}{2(1+\delta)}$. The computation can be performed numerically using $\ln(x+1) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n} x^n$, or with the help of graph software. In either way, it can be shown that the bound derived in (b) is better.

4.13

(a) From the Chernoff bound, $\Pr(X \geq xn) \leq \mathbf{E}[e^{tX}]/e^{txn}$.
Since $\mathbf{E}[e^{tX}] = \prod_{i=1}^{n} \mathbf{E}[e^{tX_i}] = (1-p+pe^t)^n$, $\Pr(X \geq xn) \leq \left(\frac{1-p+pe^t}{e^{xt}}\right)^n = ((1-p)e^{-xt} + pe^{(1-x)t})^n$. To derive the tightest bound, we solve for $\frac{\partial}{\partial t}((1-p)e^{-xt} + pe^{(1-x)t}) = 0$, which gives $t = \ln(x(1-p)) - \ln((1-x)p)$. Since $(1-p)e^{-xt} + pe^{(1-x)t}$ is convex w. r. t. t with given conditions, this gives the minimum. By plugging this in, we can show that $\Pr(X \geq xn) \leq e^{-nF(x,p)}$.

(b) Since $\frac{\partial^2}{\partial x^2}(F(x,p) - 2(x-p)^2) = \frac{1}{x} + \frac{1}{1-x} - 4 = \frac{(2x-1)^2}{x(1-x)} \geq 0$, $F(x,p) - 2(x-p)^2$ is convex w. r. t. x when $0 < x, p < 1$. Considering that $\frac{\partial}{\partial x}(F(x,p) - 2(x-p)^2) = 0$ yields $x = p$ and $(F(x,p) - 2(x-p)^2)\Big|_{x=p} = 0$, we get $F(x,p) - 2(x-p)^2 \geq 0$.
(c) $\Pr(X \geq (p+\epsilon)n) \leq e^{-nF(p+\epsilon,p)}$ holds by (a), and $e^{-nF(p+\epsilon,p)} \leq e^{-n\times 2(p+\epsilon-p)^2} = e^{-2n\epsilon^2}$ holds by (b).
(d) Take $Y_i = 1 - X_i$, and let $Y = n - X$. Then, $\Pr(X \leq (p-\epsilon)n) = \Pr(Y \geq ((1-p)+\epsilon)n) \leq e^{-2n\epsilon^2}$ holds by (c). Combined with (c), we get $\Pr(|X - pn| \geq \epsilon n) = \Pr(X \leq (p-\epsilon)n) + \Pr(X \geq (p+\epsilon)n) \leq 2e^{-2n\epsilon^2}$.

4.14

We first bound the moment generating function of X to $\mathbf{E}[e^{tX}] = \prod_{i=1}^{n}(1 - p_i + e^{ta_i}p_i) = \prod_{i=1}^{n}(1 + p_i(e^{ta_i} - 1)) \leq \prod_{i=1}^{n}\exp(p_i(e^{ta_i} - 1))$. Since $0 \leq a_i \leq 1$, $\prod_{i=1}^{n}\exp(p_i(e^{ta_i} - 1)) \leq \prod_{i=1}^{n}\exp(p_i(e^{t} - 1)) = \exp(\sum_{i=1}^{n} p_i(e^t - 1)) = \exp(\mu(e^t - 1))$.
Following the proof of Theorem 4.4 in the textbook, we get $\Pr(X \geq (1+\delta)\mu) = \Pr(e^{tX} \geq e^{t(1+\delta)\mu}) \leq \frac{\exp((e^t-1)\mu)}{\exp(t(1+\delta)\mu)}$ for $t > 0$. Now take $t = \ln(1+\delta)$, and we show the desired Chernoff bound.
Similarly, $\Pr(X \leq (1-\delta)\mu) = \Pr(e^{tX} \geq e^{t(1-\delta)\mu}) \leq \frac{\exp((e^t-1)\mu)}{\exp(t(1-\delta)\mu)}$ for $t < 0$. Now for $0 < \delta < 1$, take $t = \ln(1-\delta)$, and we can show that $\Pr(X \leq (1-\delta)\mu) \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}$.

4.15

Note that $|(1-p_i) - (-p_i)| = 1$, and $\mathbf{E}[X_i] = 0$ for all i. Applying the Hoeffding bound to X, we get $\Pr(|\frac{1}{n}\sum_{i=1}^{n} X_i| > \epsilon) = \Pr(|X| > n\epsilon) \leq 2e^{-2n\epsilon^2}$. Now take $\epsilon = \frac{a}{n}$ to get $\Pr(|X| > a) \leq 2e^{-2a^2/n}$.

4.16

Let $Y_i = a_i(X_i - p_i)$. Then we observe that $\Pr(-a_ip_i \leq Y_i \leq a_i(1-p_i)) = 1$ and $\mathbf{E}[Y_i] = 0$. Now applying the Hoeffding bound to $\sum_{i=1}^{n} Y_i$ as:
$\Pr(|\sum_{i=1}^{n} Y_i| \geq \delta\mu) \leq 2e^{-2n(\frac{\delta\mu}{n})^2\frac{1}{\max a_i^2}} \leq 2e^{-2\delta^2\mu^2/n}$.

4.17

Let the total time (in steps) of a single processor be $X = \sum_{i=1}^{n/m} X_i$, where X_i is the number of steps for the ith job of the processor. Since $\mathbf{E}[X_i] = p + (1-p)k$, we take Y_i as $X_i = 1 + (k-1)Y_i$ so that $Y_i \sim Bernoulli(1-p)$. With Y_i, we get $X = (n/m) + (k-1)\sum_{i=1}^{n/m} Y_i$.

Applying the Chernoff bounds to $Y = \sum_{i=1}^{n/m} Y_i$, we get $\Pr(|Y - \mathbf{E}[Y]| \geq \delta\mathbf{E}[Y]) \leq 2e^{-\mathbf{E}[Y]\delta^2/3}$. Since $\mathbf{E}[Y] = (n/m)(1-p)$, we can bound X as: $\Pr(|X-(n/m)(1+(1-p)(k-1))| \geq \delta(n/m)(1-p)(k-1)) \leq 2e^{-\frac{n}{m}(1-p)\frac{\delta^2}{3}}$. Using the union bound, we bound the total time T to $\Pr(|T-(n/m)(1+(1-p)(k-1))| \geq \delta(n/m)(1-p)(k-1)) \leq 2me^{-\frac{n}{m}(1-p)\frac{\delta^2}{3}}$. We can take $\delta = \sqrt{\frac{3\ln(2m/\epsilon)}{(n/m)(1-p)}}$ to derive the bound with probability of at least $1-\epsilon$.

4.18