3 Feb 2025 — The Chernoff Bound

Announcements.

1. Quiz 1 grades on Gradescope. Regrades until Sunday.
2. PSet 1 on Canvas. Turn in on Gradescope.
   Deadline Tues 2/11. Late deadline noon on 2/13.

Q. When is the earliest $m(n)$ s.t. with probability at least $1/2$,

$$\frac{\text{max occupancy}}{\text{min occupancy}} \leq 1+\varepsilon?$$

Method. Tail bounds and union bounds

Look at occupancy on one bin, i, in isolation.

$$\mathbb{E}[\#\text{ balls in bin } i] = \frac{m}{n}$$

If $m \gg n$ the "law of large numbers" (don't worry about what that is) says actual # balls in bin i is very unlikely to be far from its exp. val:

Look for $\delta > 0$ such that

$$\frac{(1+\delta)\,m/n}{(1-\delta)\,m/n} \leq 1+\varepsilon.$$

GOOD NEWS! $\delta = \varepsilon/3$ works.

Now focus on showing that with prob $\geq \frac{1}{2}$, every bin's occupancy, L_i, satisfies

$$\boxed{\forall i \quad (1-\delta)\frac{m}{n} \le L_i \le (1+\delta)\frac{m}{n}.}$$

Then it'll follow that

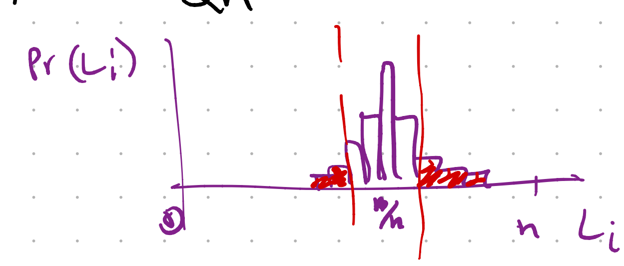
$$(1-\delta)\frac{m}{n} \le \text{min occ.} \le \text{Max occ.} \le (1+\delta)\frac{m}{n}$$

So

$$\frac{\max}{\min} \le \frac{1+\delta}{1-\delta} \le 1+\varepsilon \quad \text{.. by choice of } \delta.$$

Now, search for m large enough to justify

$$\text{(TB)} \quad \forall i \quad \Pr\left(\left|L_i - \frac{m}{n}\right| > \frac{\delta m}{n}\right) \le \frac{1}{2n}$$



To finish up, use the union bound

$$\text{(UB)} \quad \Pr(\mathcal{E}_1 \cup \mathcal{E}_2 \cup \dots \cup \mathcal{E}_N) \le \sum_{j=1}^{N} \Pr(\mathcal{E}_j).$$

With $\mathcal{E}_{i,\delta}$ denoting the event $\left\{ \left|L_i - \frac{m}{n}\right| > \frac{\delta m}{n} \right\}$

$$\text{(TB)} + \text{(UB)} \Longrightarrow \Pr(\exists i \ \mathcal{E}_{i,\delta}) \le \sum_{i=1}^{n} \Pr(\mathcal{E}_{i,\delta}) \le n \cdot \left(\frac{1}{2n}\right) = \frac{1}{2}.$$

So, $\Pr\left(\boxed{\forall i \ \overline{\mathcal{E}_{i,\delta}}}\right) \ge \frac{1}{2}.$

How large does m need to be for TB to hold?

$$\text{Chebyshev} \Longrightarrow O(n^2/\varepsilon^2)$$

$$\text{Chernoff} \Longrightarrow O(n (\log n)/\varepsilon^2)$$

Recall: Chebyshev says a random variable is unlikely to be far from its expectation if variance is small.

$$\Pr(|X - \mathbb{E}[X]| \geq \lambda) \leq \frac{Var(X)}{\lambda^2}$$

For occupancy of bin i, L_i, it is a sum of m independent Bernoulli random vars ($\{0,1\}$-valued) each with expected value $\frac{1}{n}$.

For a Bernoulli Y with exp val p,

$$\mathbb{E}[Y^2] = \mathbb{E}[Y] = p.$$

$$Var(Y) = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 = p - p^2 = p(1-p).$$

$$Var(L_i) = m \cdot Var(Y)$$ where Y is Bernoulli$(\frac{1}{n})$

$$= m \cdot \frac{1}{n} \cdot \left(1 - \frac{1}{n}\right) < \frac{m}{n}.$$

$$\Pr\left(\left|L_i - \frac{m}{n}\right| \geq \frac{\delta m}{n}\right) \leq \frac{\mathrm{Var}(L_i)}{(\delta m/n)^2}$$

$$< \frac{m/n}{\delta^2 \cdot (m/n)^2} = \frac{n}{\delta^2 m} = \frac{9n}{\varepsilon^2 m}.$$

Remember: we were trying to choose m large enough that this probability is $\leq \frac{1}{2n}$.

$$m \geq 18n^2/\varepsilon^2$$

$$\frac{9n}{\varepsilon^2 m} \leq \frac{9n}{18n^2} = \frac{1}{2n},$$

The Moment Generating Function

If X is a random variable its moment generating function is

$$M_X(t) = \mathbb{E}\left[e^{tX}\right]$$

The <u>cumulant generating function</u> is

$$K_X(t) = \ln M_X(t)$$

<u>Lemma</u>, If $X_1, \dots, X_N$ are independent random variables and $X = X_1 + \dots + X_N$ then

$$M_X(t) = \prod_{i=1}^{N} M_{X_i}(t)$$

$$K_X(t) = \sum_{i=1}^{N} K_{X_i}(t)$$

... when the LHS and RHS are finite.

<u>Proof.</u> $M_X(t) = \mathbb{E}[e^{tX}]$

$$= \mathbb{E}\left[e^{t(X_1 + \dots + X_N)}\right]$$

$$= \mathbb{E}\left[\prod_{i=1}^{N} e^{tX_i}\right] = \prod_{i=1}^{N}\left[\mathbb{E}[e^{tX_i}]\right]$$

$$= \prod_{i=1}^{N} M_{X_i}(t)$$

Taking $\ln(\cdot)$ of both sides,

$$K_X(t) = \sum_{i=1}^{N} K_{X_i}(t)$$

<u>Lemma</u> If $Y_1 \ldots Y_N$ mutually independent then

$$\mathbb{E}\left(\prod_{i=1}^{N} Y_i\right) = \prod_{i=1}^{N} \mathbb{E}[Y_i]$$

<u>Pf.</u> If $Y_1, \ldots, Y_N$ are independent

$$\mathbb{E}\left[\prod_{i=1}^{N} Y_i\right] = \sum_{\vec{y} = (y_1, \ldots, y_N)} \Pr(\forall i\ Y_i = y_i) \cdot \left(\prod_{i=1}^{N} y_i\right)$$

$$= \sum_{\vec{y}} \prod_{i=1}^{N} \Pr(Y_i = y_i) \cdot \prod_{i=1}^{N} y_i$$

$$= \sum_{\vec{y}} \prod_{i=1}^{N} \left(\Pr(Y_i = y_i) \cdot y_i\right)$$

$$= \prod_{i=1}^{N} \left(\sum_{y_i} \Pr(Y_i = y_i) \cdot y_i\right)$$

$$= \prod_{i=1}^{N} \mathbb{E}[Y_i]$$

The Chernoff Bound.

If $X_1, X_2, \ldots, X_N$ are mutually independent random variables taking values in $[0,1]$ and $X = X_1 + \ldots + X_N$ then $\forall\ 0 < \varepsilon < 1$

$$\Pr\left(X > (1+\varepsilon) \cdot \mathbb{E}[X] \right) < e^{-\frac{1}{3}\varepsilon^2 \mathbb{E}[X]}$$

$$\Pr\left(X < (1-\varepsilon)\, \mathbb{E}[X] \right) < e^{-\frac{1}{2}\varepsilon^2 \mathbb{E}[X]}$$