Randomized Algorithms, Lecture 4

Jacob Holm (jaho@di.ku.dk)

May 6th 2019

Good Morning.

Today's Lecture

(More) Tail Inequalities

- The Coupon Collectors Problem
- The Chernoff Bound

Last time I talked about Occupancy Problems, proved Markov's and Chebyshev's Tail Inequalities and showed an example (LazySelect) of their use.
Today, we will see one more abstract problem in the form of the *Coupon collecters problem*.
Then I'll present a technique known as the Chernoff bound, which for many problems allows us to get *exponentially small* probability bounds, instead of the polynomially small bounds we usually get from Chebyshev.

Markov and Chebyshev

Theorem

Let Y be a random variable taking only non-negative values. Then for all $t, k > 0$:

$$\Pr[Y \geq t] \leq \frac{\mathbb{E}[Y]}{t} \quad \wedge \quad \Pr[Y \geq k\,\mathbb{E}[Y]] \leq \frac{1}{k}$$

Theorem

Let $X \in \mathbb{R}$ be a random variable with expectation $\mu_X = \mathbb{E}[X]$ and standard deviation $\sigma_X = \sqrt{\mathbb{E}[(X - \mu_X)^2]}$. Then for all $t > 0$:

$$\Pr[|X - \mu_X| \geq t\sigma_X] \leq \frac{1}{t^2}$$

Recall: Given a random variable X with expectation $\mathbb{E}[X] = \mu_X$, define its *variance* as $\sigma_X^2 := \mathbb{E}[(X - \mu_X)^2]$, and its *standard deviation* as $\sigma_X := \sqrt{\mathbb{E}[(X - \mu_X)^2]}$.

The Coupon Collectors Problem

Suppose there are n types of coupons, and at each trial a coupon is chosen at random. Each random coupon is equally likely to be of any of the n types, and the choices are independent.

How many trials do we expect before we have at least one of each type of coupon? And how likely is it that the number of trials deviate significantly from its expectation?

The Coupon Collectors Problem

Suppose there are n types of coupons, and at each trial a coupon is chosen at random. Each random coupon is equally likely to be of any of the n types, and the choices are independent.
How many trials do we expect before we have at least one of each type of coupon?
And how likely is it that the number of trials deviate significantly from its expectation?

Coupon Collector, Analysis.

Let X be the number of trials before we have at least one of each coupon type.

Let $C_1, \ldots, C_X \in \{1, \ldots, n\}$ be the result of each trial. Trial C_i is a *success* if $C_i \notin \{C_1, \ldots, C_{i-1}\}$.

Define *epoch* $i \in \{0, \ldots, n-1\}$ to consist of all trials from just after the ith success, until the $(i+1)$st success. Let X_i be the number of trials in epoch i, so $X = \sum_{i=0}^{n-1} X_i$

Let p_i be the probability of success in epoch i.

What is p_i?

Coupon Collector, Analysis.

Let X be the number of trials before we have at least one of each coupon type.

Let $C_1, \ldots, C_X \in \{1, \ldots, n\}$ be the result of each trial. Trial C_i is a *success* if $C_i \notin \{C_1, \ldots, C_{i-1}\}$.

Define *epoch* $i \in \{0, \ldots, n-1\}$ to consist of all trials from just after the ith success, until the $(i+1)$st success. Let X_i be the number of trials in epoch i, so $X = \sum_{i=0}^{n-1} X_i$

Let p_i be the probability of success in epoch i.

What is p_i?

Coupon Collector, Analysis.

Let X be the number of trials before we have at least one of each coupon type.

Let $C_1, \ldots, C_X \in \{1, \ldots, n\}$ be the result of each trial. Trial C_i is a *success* if $C_i \notin \{C_1, \ldots, C_{i-1}\}$.

Define *epoch* $i \in \{0, \ldots, n-1\}$ to consist of all trials from just after the ith success, until the $(i+1)$st success. Let X_i be the number of trials in epoch i, so $X = \sum_{i=0}^{n-1} X_i$

Let p_i be the probability of success in epoch i.

What is p_i?

Coupon Collector, Analysis.

Let X be the number of trials before we have at least one of each coupon type.

Let $C_1, \ldots, C_X \in \{1, \ldots, n\}$ be the result of each trial. Trial C_i is a *success* if $C_i \notin \{C_1, \ldots, C_{i-1}\}$.

Define *epoch* $i \in \{0, \ldots, n-1\}$ to consist of all trials from just after the ith success, until the $(i+1)$st success. Let X_i be the number of trials in epoch i, so $X = \sum_{i=0}^{n-1} X_i$

Let p_i be the probability of success in epoch i.

What is p_i?

Coupon Collector, Analysis.

Let X be the number of trials before we have at least one of each coupon type.

Let $C_1, \ldots, C_X \in \{1, \ldots, n\}$ be the result of each trial. Trial C_i is a *success* if $C_i \notin \{C_1, \ldots, C_{i-1}\}$.

Define *epoch* $i \in \{0, \ldots, n-1\}$ to consist of all trials from just after the ith success, until the $(i+1)$st success. Let X_i be the number of trials in epoch i, so $X = \sum_{i=0}^{n-1} X_i$

Let p_i be the probability of success in epoch i.

What is p_i? $p_i = \frac{n-i}{n}$

Coupon Collector, Analysis.

Each X_i is *geometrically* distributed, so $\mathbb{E}[X_i] = \frac{1}{p_i}$ and $\sigma^2_{X_i} = \frac{1-p_i}{p_i^2}$. Thus

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=0}^{n-1} X_i\right] = \sum_{i=0}^{n-1} \mathbb{E}[X_i] = \sum_{i=0}^{n-1} \frac{1}{p_i}$$

$$= \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^{n} \frac{1}{j} = nH_n = n \ln n + \mathcal{O}(n)$$

$$\sigma^2_X =$$

Coupon Collector, Analysis.

Each X_i is *geometrically* distributed, so $\mathbb{E}[X_i] = \frac{1}{p_i}$ and $\sigma^2_{X_i} = \frac{1-p_i}{p_i^2}$. Thus

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=0}^{n-1} X_i\right] = \sum_{i=0}^{n-1} \mathbb{E}[X_i] = \sum_{i=0}^{n-1} \frac{1}{p_i}$$

$$= \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^{n} \frac{1}{j} = nH_n = n \ln n + \mathcal{O}(n)$$

$$\sigma^2_X =$$

Coupon Collector, Analysis.

Each X_i is *geometrically* distributed, so $\mathbb{E}[X_i] = \frac{1}{p_i}$ and $\sigma^2_{X_i} = \frac{1-p_i}{p_i^2}$. Thus

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=0}^{n-1} X_i\right] = \sum_{i=0}^{n-1} \mathbb{E}[X_i] = \sum_{i=0}^{n-1} \frac{1}{p_i}$$

$$= \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^{n} \frac{1}{j} = nH_n = n \ln n + \mathcal{O}(n)$$

$$\sigma^2_X =$$

Coupon Collector, Analysis.

Each X_i is *geometrically* distributed, so $\mathbb{E}[X_i] = \frac{1}{p_i}$ and $\sigma^2_{X_i} = \frac{1-p_i}{p_i^2}$. Thus

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=0}^{n-1} X_i\right] = \sum_{i=0}^{n-1} \mathbb{E}[X_i] = \sum_{i=0}^{n-1} \frac{1}{p_i}$$

$$= \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^{n} \frac{1}{j} = nH_n = n \ln n + \mathcal{O}(n)$$

$$\sigma^2_X =$$

Coupon Collector, Analysis.

Each X_i is *geometrically* distributed, so $\mathbb{E}[X_i] = \frac{1}{p_i}$ and $\sigma^2_{X_i} = \frac{1-p_i}{p_i^2}$. Thus

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=0}^{n-1} X_i\right] = \sum_{i=0}^{n-1} \mathbb{E}[X_i] = \sum_{i=0}^{n-1} \frac{1}{p_i}$$

$$= \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^{n} \frac{1}{j} = nH_n = n \ln n + \mathcal{O}(n)$$

$$\sigma^2_X =$$

Coupon Collector, Analysis.

Each X_i is *geometrically* distributed, so $\mathbb{E}[X_i] = \frac{1}{p_i}$ and $\sigma^2_{X_i} = \frac{1-p_i}{p_i^2}$. Thus

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=0}^{n-1} X_i\right] = \sum_{i=0}^{n-1} \mathbb{E}[X_i] = \sum_{i=0}^{n-1} \frac{1}{p_i}$$

$$= \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^{n} \frac{1}{j} = nH_n = n \ln n + \mathcal{O}(n)$$

$$\sigma^2_X =$$

Coupon Collector, Analysis.

Each X_i is *geometrically* distributed, so $\mathbb{E}[X_i] = \frac{1}{p_i}$ and $\sigma^2_{X_i} = \frac{1-p_i}{p_i^2}$. Thus

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=0}^{n-1} X_i\right] = \sum_{i=0}^{n-1} \mathbb{E}[X_i] = \sum_{i=0}^{n-1} \frac{1}{p_i}$$
$$= \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^{n} \frac{1}{j} = nH_n = n \ln n + \mathcal{O}(n)$$

$$\sigma^2_X =$$

Coupon Collector, Analysis.

Each X_i is *geometrically* distributed, so $\mathbb{E}[X_i] = \frac{1}{p_i}$ and $\sigma_{X_i}^2 = \frac{1-p_i}{p_i^2}$. Thus

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}\Big[\sum_{i=0}^{n-1} X_i\Big] = \sum_{i=0}^{n-1} \mathbb{E}[X_i] = \sum_{i=0}^{n-1} \frac{1}{p_i} \\
&= \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^{n} \frac{1}{j} = nH_n = n \ln n + \mathcal{O}(n)
\end{aligned}$$

$$\sigma_X^2 =$$

Coupon Collector, Analysis.

Each X_i is *geometrically* distributed, so $\mathbb{E}[X_i] = \frac{1}{p_i}$ and $\sigma^2_{X_i} = \frac{1-p_i}{p_i^2}$. Thus

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=0}^{n-1} X_i\right] = \sum_{i=0}^{n-1} \mathbb{E}[X_i] = \sum_{i=0}^{n-1} \frac{1}{p_i}$$

$$= \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^{n} \frac{1}{j} = nH_n = n \ln n + \mathcal{O}(n)$$

$$\sigma^2_X = \sum_{i=0}^{n-1} \sigma^2_{X_i} = \sum_{i=0}^{n-1} \frac{ni}{(n-i)^2} = \sum_{j=1}^{n} \frac{n(n-j)}{j^2}$$

$$= n^2 \sum_{j=1}^{n} \frac{1}{j^2} - n \sum_{j=1}^{n} \frac{1}{j} \leq n^2 \frac{\pi^2}{6} - nH_n$$

$$\sigma_X < \frac{\pi n}{\sqrt{6}}$$

Coupon Collector, Analysis.

Each X_i is *geometrically* distributed, so $\mathbb{E}[X_i] = \frac{1}{p_i}$ and $\sigma^2_{X_i} = \frac{1-p_i}{p_i^2}$. Thus

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=0}^{n-1} X_i\right] = \sum_{i=0}^{n-1} \mathbb{E}[X_i] = \sum_{i=0}^{n-1} \frac{1}{p_i}$$

$$= \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^{n} \frac{1}{j} = nH_n = n \ln n + \mathcal{O}(n)$$

$$\sigma_X^2 = \sum_{i=0}^{n-1} \sigma^2_{X_i} = \sum_{i=0}^{n-1} \frac{ni}{(n-i)^2} = \sum_{j=1}^{n} \frac{n(n-j)}{j^2}$$

$$= n^2 \sum_{j=1}^{n} \frac{1}{j^2} - n \sum_{j=1}^{n} \frac{1}{j} \leq n^2 \frac{\pi^2}{6} - nH_n$$

$$\sigma_X < \frac{\pi n}{\sqrt{6}}$$

$$\frac{1 - \frac{n-i}{n}}{\left(\frac{n-i}{n}\right)^2} = \frac{\frac{i}{n}}{\left(\frac{n-i}{n}\right)^2} = \frac{ni}{(n-i)^2}$$

Coupon Collector, Analysis.

Each X_i is *geometrically* distributed, so $\mathbb{E}[X_i] = \frac{1}{p_i}$ and $\sigma_{X_i}^2 = \frac{1-p_i}{p_i^2}$. Thus

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=0}^{n-1} X_i\right] = \sum_{i=0}^{n-1} \mathbb{E}[X_i] = \sum_{i=0}^{n-1} \frac{1}{p_i}$$

$$= \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^{n} \frac{1}{j} = nH_n = n \ln n + \mathcal{O}(n)$$

$$\sigma_X^2 = \sum_{i=0}^{n-1} \sigma_{X_i}^2 = \sum_{i=0}^{n-1} \frac{ni}{(n-i)^2} = \sum_{j=1}^{n} \frac{n(n-j)}{j^2}$$

$$= n^2 \sum_{j=1}^{n} \frac{1}{j^2} - n \sum_{j=1}^{n} \frac{1}{j} \leq n^2 \frac{\pi^2}{6} - nH_n$$

$$\sigma_X < \frac{\pi n}{\sqrt{6}}$$

Coupon Collector, Analysis.

Each X_i is *geometrically* distributed, so $\mathbb{E}[X_i] = \frac{1}{p_i}$ and $\sigma^2_{X_i} = \frac{1-p_i}{p_i^2}$. Thus

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=0}^{n-1} X_i\right] = \sum_{i=0}^{n-1} \mathbb{E}[X_i] = \sum_{i=0}^{n-1} \frac{1}{p_i}$$

$$= \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^{n} \frac{1}{j} = nH_n = n \ln n + \mathcal{O}(n)$$

$$\sigma^2_X = \sum_{i=0}^{n-1} \sigma^2_{X_i} = \sum_{i=0}^{n-1} \frac{ni}{(n-i)^2} = \sum_{j=1}^{n} \frac{n(n-j)}{j^2}$$

$$= n^2 \sum_{j=1}^{n} \frac{1}{j^2} - n \sum_{j=1}^{n} \frac{1}{j} \leq n^2 \frac{\pi^2}{6} - nH_n$$

$$\sigma_X < \frac{\pi n}{\sqrt{6}}$$

Coupon Collector, Analysis.

Each X_i is *geometrically* distributed, so $\mathbb{E}[X_i] = \frac{1}{p_i}$ and $\sigma^2_{X_i} = \frac{1-p_i}{p_i^2}$. Thus

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}\left[\sum_{i=0}^{n-1} X_i\right] = \sum_{i=0}^{n-1} \mathbb{E}[X_i] = \sum_{i=0}^{n-1} \frac{1}{p_i} \\
&= \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^{n} \frac{1}{j} = nH_n = n \ln n + \mathcal{O}(n) \\
\sigma^2_X &= \sum_{i=0}^{n-1} \sigma^2_{X_i} = \sum_{i=0}^{n-1} \frac{ni}{(n-i)^2} = \sum_{j=1}^{n} \frac{n(n-j)}{j^2} \\
&= n^2 \sum_{j=1}^{n} \frac{1}{j^2} - n \sum_{j=1}^{n} \frac{1}{j} \leq n^2 \frac{\pi^2}{6} - nH_n
\end{aligned}$$

$$\sigma_X < \frac{\pi n}{\sqrt{6}}$$

Coupon Collector, Analysis.

Each X_i is *geometrically* distributed, so $\mathbb{E}[X_i] = \frac{1}{p_i}$ and $\sigma^2_{X_i} = \frac{1-p_i}{p_i^2}$. Thus

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=0}^{n-1} X_i\right] = \sum_{i=0}^{n-1} \mathbb{E}[X_i] = \sum_{i=0}^{n-1} \frac{1}{p_i}$$

$$= \sum_{i=0}^{n-1} \frac{n}{n-i} = n \sum_{j=1}^{n} \frac{1}{j} = nH_n = n \ln n + \mathcal{O}(n)$$

$$\sigma^2_X = \sum_{i=0}^{n-1} \sigma^2_{X_i} = \sum_{i=0}^{n-1} \frac{ni}{(n-i)^2} = \sum_{j=1}^{n} \frac{n(n-j)}{j^2}$$

$$= n^2 \sum_{j=1}^{n} \frac{1}{j^2} - n \sum_{j=1}^{n} \frac{1}{j} \leq n^2 \frac{\pi^2}{6} - nH_n$$

$$\sigma_X < \frac{\pi n}{\sqrt{6}}$$

Coupon Collector, Analysis.

Trying Chebyshev gives (for $\beta > 1$)

$$\begin{aligned}
\Pr[X > \beta n \ln n] &\le \Pr[|X - \mu_X| \ge (\beta - 1) n \ln n] \\
&\le \Pr[|X - \mu_X| \ge (\beta - 1)\left(\tfrac{\sqrt{6}}{\pi}\sigma_X\right) \ln n] \\
&\le \frac{\pi}{6(\beta - 1)^2 \ln^2 n} = \mathcal{O}\left(\frac{1}{\ln^2 n}\right)
\end{aligned}$$

Is this a good bound?

Coupon Collector, Analysis.

Trying Chebyshev gives (for $\beta > 1$)

$$\begin{aligned}
\Pr[X > \beta n \ln n] &\le \Pr[|X - \mu_X| \ge (\beta - 1) n \ln n] \\
&\le \Pr[|X - \mu_X| \ge (\beta - 1)\left(\tfrac{\sqrt{6}}{\pi}\sigma_X\right) \ln n] \\
&\le \frac{\pi}{6(\beta - 1)^2 \ln^2 n} = \mathcal{O}\left(\frac{1}{\ln^2 n}\right)
\end{aligned}$$

Is this a good bound?

Coupon Collector, Analysis.

Trying Chebyshev gives (for $\beta > 1$)

$$\begin{aligned}
\Pr[X > \beta n \ln n] &\le \Pr[|X - \mu_X| \ge (\beta - 1) n \ln n] \\
&\le \Pr[|X - \mu_X| \ge (\beta - 1)(\tfrac{\sqrt{6}}{\pi}\sigma_X) \ln n] \\
&\le \frac{\pi}{6(\beta - 1)^2 \ln^2 n} = \mathcal{O}\left(\frac{1}{\ln^2 n}\right)
\end{aligned}$$

Is this a good bound? NO!

Coupon Collector, Analysis.

Let $\mathcal{E}_i^r$ be the event that coupon i is *not* collected in the first r trials. Then

$$\Pr[\mathcal{E}_i^r] = \left(1 - \frac{1}{n}\right)^r \leq$$

Since the probability of a union of events is at most the sum of probabilities, we have for $r = \beta n \ln n$

$$\Pr[X > r] = \Pr[\cup_{i=1}^n \mathcal{E}_i^r] \leq \sum_{i=1}^n \Pr[\mathcal{E}_i^r]$$

$$\leq \sum_{i=1}^n e^{-\frac{r}{n}} = \sum_{i=1}^n n^{-\beta} = n^{-(\beta-1)}$$

Coupon Collector, Analysis.

Let $\mathcal{E}_i^r$ be the event that coupon i is *not* collected in the first r trials. Then

$$\Pr[\mathcal{E}_i^r] = \left(1 - \tfrac{1}{n}\right)^r \le e^{-\frac{r}{n}} \qquad \text{(Using } 1 + x \le e^x\text{)}$$

Since the probability of a union of events is at most the sum of probabilities, we have for $r = \beta n \ln n$

$$\Pr[X > r] = \Pr[\cup_{i=1}^n \mathcal{E}_i^r] \le \sum_{i=1}^n \Pr[\mathcal{E}_i^r]$$
$$\le \sum_{i=1}^n e^{-\frac{r}{n}} = \sum_{i=1}^n n^{-\beta} = n^{-(\beta-1)}$$

Coupon Collector, Analysis.

Let $\mathcal{E}_i^r$ be the event that coupon i is *not* collected in the first r trials. Then

$$\Pr[\mathcal{E}_i^r] = \left(1 - \tfrac{1}{n}\right)^r \leq e^{-\frac{r}{n}} \qquad \text{(Using } 1 + x \leq e^x\text{)}$$

Since the probability of a union of events is at most the sum of probabilities, we have for $r = \beta n \ln n$

$$\Pr[X > r] = \Pr[\cup_{i=1}^n \mathcal{E}_i^r] \leq \sum_{i=1}^{n} \Pr[\mathcal{E}_i^r]$$

$$\leq \sum_{i=1}^{n} e^{-\frac{r}{n}} = \sum_{i=1}^{n} n^{-\beta} = n^{-(\beta-1)}$$

Coupon Collector, Analysis.

Let $\mathcal{E}_i^r$ be the event that coupon i is *not* collected in the first r trials. Then

$$\Pr[\mathcal{E}_i^r] = \left(1 - \tfrac{1}{n}\right)^r \leq e^{-\frac{r}{n}} \qquad (\text{Using } 1 + x \leq e^x)$$

Since the probability of a union of events is at most the sum of probabilities, we have for $r = \beta n \ln n$

$$\Pr[X > r] = \Pr[\cup_{i=1}^n \mathcal{E}_i^r] \leq \sum_{i=1}^n \Pr[\mathcal{E}_i^r]$$

$$\leq \sum_{i=1}^n e^{-\frac{r}{n}} = \sum_{i=1}^n n^{-\beta} = n^{-(\beta-1)}$$

Coupon Collector, Analysis.

Let $\mathcal{E}_i^r$ be the event that coupon i is *not* collected in the first r trials. Then

$$\Pr[\mathcal{E}_i^r] = \left(1 - \frac{1}{n}\right)^r \leq e^{-\frac{r}{n}} \qquad (\text{Using } 1 + x \leq e^x)$$

Since the probability of a union of events is at most the sum of probabilities, we have for $r = \beta n \ln n$

$$\Pr[X > r] = \Pr[\cup_{i=1}^n \mathcal{E}_i^r] \leq \sum_{i=1}^n \Pr[\mathcal{E}_i^r]$$

$$\leq \sum_{i=1}^n e^{-\frac{r}{n}} = \sum_{i=1}^n n^{-\beta} = n^{-(\beta-1)}$$

Coupon Collector, Analysis.

Let $\mathcal{E}_i^r$ be the event that coupon i is *not* collected in the first r trials. Then

$$\Pr[\mathcal{E}_i^r] = \left(1 - \tfrac{1}{n}\right)^r \leq e^{-\frac{r}{n}} \qquad (\text{Using } 1 + x \leq e^x)$$

Since the probability of a union of events is at most the sum of probabilities, we have for $r = \beta n \ln n$

$$\begin{aligned}\Pr[X > r] = \Pr[\cup_{i=1}^n \mathcal{E}_i^r] &\leq \sum_{i=1}^n \Pr[\mathcal{E}_i^r] \\ &\leq \sum_{i=1}^n e^{-\frac{r}{n}} = \sum_{i=1}^n n^{-\beta} = n^{-(\beta-1)}\end{aligned}$$

Coupon Collector, Summary.

- The expected number of trials is $\mathbb{E}[X] = n \ln n + \mathcal{O}(n)$.
- The probability of using more than $r = \beta n \ln n$ trials for some $\beta > 1$ is at most $\Pr[X > r] \leq n^{-(\beta-1)}$.
- This does *not* follow from a simple application of Chebyshev.
- A more complex proof shows that for $c \in \mathbb{R}^+$, $\lim_{n\to\infty} \Pr[|X - n \ln n| \leq cn] = e^{-e^{-c}} - e^{-e^{c}}$

Coupon Collector, Summary.

- The expected number of trials is $\mathbb{E}[X] = n \ln n + \mathcal{O}(n)$.
- The probability of using more than $r = \beta n \ln n$ trials for some $\beta > 1$ is at most $\Pr[X > r] \leq n^{-(\beta-1)}$.
- This does *not* follow from a simple application of Chebyshev.
- A more complex proof shows that for $c \in \mathbb{R}^+$, $\lim_{n\to\infty} \Pr[|X - n \ln n| \leq cn] = e^{-e^{-c}} - e^{-e^{c}}$

Coupon Collector, Summary.

- The expected number of trials is $\mathbb{E}[X] = n \ln n + \mathcal{O}(n)$.
- The probability of using more than $r = \beta n \ln n$ trials for some $\beta > 1$ is at most $\Pr[X > r] \leq n^{-(\beta-1)}$.
- This does *not* follow from a simple application of Chebyshev.
- A more complex proof shows that for $c \in \mathbb{R}^+$, $\lim_{n\to\infty} \Pr[|X - n \ln n| \leq cn] = e^{-e^{-c}} - e^{-e^{c}}$

Coupon Collector, Summary.

- The expected number of trials is $\mathbb{E}[X] = n \ln n + \mathcal{O}(n)$.
- The probability of using more than $r = \beta n \ln n$ trials for some $\beta > 1$ is at most $\Pr[X > r] \leq n^{-(\beta-1)}$.
- This does *not* follow from a simple application of Chebyshev.
- A more complex proof shows that for $c \in \mathbb{R}^+$, $\lim_{n\to\infty} \Pr[|X - n \ln n| \leq cn] = e^{-e^{-c}} - e^{-e^{c}}$

Bernoulli trials

Think of this as a sequence of coin flips, with the same (possibly biased) coin.

Let $0 \leq p \leq 1$, let $X_1, \ldots, X_n$ be independent indicator variables with $\Pr[X_i = 1] = p$, and let $X = \sum_{i=1}^{n} X_i$.

We call $X_1, \ldots X_n$ *Bernoulli Trials*, and say that X has the *Binomial Distribution*.

Poisson trials

Let $0 \leq p_1, \ldots, p_n \leq 1$, let $X_1, \ldots, X_n$ be independent indicator variables with $\Pr[X_i = 1] = p_i$, and let $X = \sum_{i=1}^{n} X_i$.

We call $X_1, \ldots X_n$ *Poisson Trials*, and say that X has the *Poisson Binomial Distribution*.

Not to be confused with the *Poisson Distribution*. This can again be thought of as a sequence of coin flips, but this time each flip is of a different coin, each coin with its own bias.

Tail Inequalities (again)

Given a (Poisson) Binomially Distributed random variable X:

- For $\delta > 0$, find $\epsilon > 0$ so

$$\Pr[X > (1 + \delta)\mu] < \epsilon$$

- For $\epsilon > 0$, find $\delta > 0$ so

$$\Pr[X > (1 + \delta)\mu] < \epsilon$$

- For $0 < \delta < 1$, find $\epsilon > 0$ so

$$\Pr[X < (1 - \delta)\mu] < \epsilon$$

- For $\epsilon > 0$, find $0 < \delta < 1$ so

$$\Pr[X < (1 - \delta)\mu] < \epsilon$$

Tail Inequalities (again)

Given a (Poisson) Binomially Distributed random variable X:

- For $\delta > 0$, find $\epsilon > 0$ so
$$\Pr[X > (1+\delta)\mu] < \epsilon$$
- For $\epsilon > 0$, find $\delta > 0$ so
$$\Pr[X > (1+\delta)\mu] < \epsilon$$
- For $0 < \delta < 1$, find $\epsilon > 0$ so
$$\Pr[X < (1-\delta)\mu] < \epsilon$$
- For $\epsilon > 0$, find $0 < \delta < 1$ so
$$\Pr[X < (1-\delta)\mu] < \epsilon$$

Tail Inequalities (again)

Given a (Poisson) Binomially Distributed random variable X:

- For $\delta > 0$, find $\epsilon > 0$ so

$$\Pr[X > (1+\delta)\mu] < \epsilon$$

- For $\epsilon > 0$, find $\delta > 0$ so

$$\Pr[X > (1+\delta)\mu] < \epsilon$$

- For $0 < \delta < 1$, find $\epsilon > 0$ so

$$\Pr[X < (1-\delta)\mu] < \epsilon$$

- For $\epsilon > 0$, find $0 < \delta < 1$ so

$$\Pr[X < (1-\delta)\mu] < \epsilon$$

Tail Inequalities (again)

Given a (Poisson) Binomially Distributed random variable X:

- For $\delta > 0$, find $\epsilon > 0$ so
$$\Pr[X > (1+\delta)\mu] < \epsilon$$
- For $\epsilon > 0$, find $\delta > 0$ so
$$\Pr[X > (1+\delta)\mu] < \epsilon$$
- For $0 < \delta < 1$, find $\epsilon > 0$ so
$$\Pr[X < (1-\delta)\mu] < \epsilon$$
- For $\epsilon > 0$, find $0 < \delta < 1$ so
$$\Pr[X < (1-\delta)\mu] < \epsilon$$

Tail Inequalities (again)

Given a (Poisson) Binomially Distributed random variable X:

- For $\delta > 0$, find $\epsilon > 0$ so
$$\Pr[X > (1+\delta)\mu] < \epsilon$$
- For $\epsilon > 0$, find $\delta > 0$ so
$$\Pr[X > (1+\delta)\mu] < \epsilon$$
- For $0 < \delta < 1$, find $\epsilon > 0$ so
$$\Pr[X < (1-\delta)\mu] < \epsilon$$
- For $\epsilon > 0$, find $0 < \delta < 1$ so
$$\Pr[X < (1-\delta)\mu] < \epsilon$$

First Chernoff Bound

Theorem

Let $X_1, \ldots, X_n$ be independent Poisson trials such that, for $1 \leq i \leq n$, $\Pr[X_i = 1] = p_i$, where $0 < p_i < 1$. Then, for $X = \sum_{i=1}^n X_i$, $\mu = \mathbb{E}[X] = \sum_{i=1}^n p_i$, and any $\delta > 0$.

$$\Pr[X > (1+\delta)\mu] < \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$$

Chernoff Bound, Main Ideas

- Analyze e^{tX} (for some $t > 0$) rather than X.
- Use independence to turn expectation of a product into a product of expectations.
- Pick t to get best possible bound.

First Chernoff Bound, Proof

For any $t > 0$:

$$
\begin{aligned}
\Pr[X > (1+\delta)\mu] &= \Pr[e^{tX} > e^{t(1+\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^{n} \mathbb{E}\left[e^{tX_i}\right]}{e^{t(1+\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^{n}(1 + p_i(e^t - 1))}{e^{t(1+\delta)\mu}} && \text{(By def of } \mathbb{E}) \\
&\le \frac{\prod_{i=1}^{n} e^{p_i(e^t-1)}}{e^{t(1+\delta)\mu}} && \text{(By } 1 + x \le e^x) \\
&= \frac{e^{\left(\sum_{i=1}^{n} p_i(e^t-1)\right)}}{e^{t(1+\delta)\mu}} \\
&= \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} = \left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^{\mu}
\end{aligned}
$$

First Chernoff Bound, Proof

For any $t > 0$:

$$\begin{aligned}
\Pr[X > (1+\delta)\mu] &= \Pr[e^{tX} > e^{t(1+\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^{n} \mathbb{E}\left[e^{tX_i}\right]}{e^{t(1+\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^{n}(1 + p_i(e^t - 1))}{e^{t(1+\delta)\mu}} && \text{(By def of } \mathbb{E}) \\
&\le \frac{\prod_{i=1}^{n} e^{p_i(e^t-1)}}{e^{t(1+\delta)\mu}} && \text{(By } 1 + x \le e^x) \\
&= \frac{e^{\left(\sum_{i=1}^{n} p_i(e^t-1)\right)}}{e^{t(1+\delta)\mu}} \\
&= \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} = \left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^{\mu}
\end{aligned}$$

Strict, because $0 < p_i < 1$ is strict.

First Chernoff Bound, Proof

$$\mathbb{E}[e^{tX}] = \mathbb{E}[e^{t\sum_{i=1}^n X_i}] = \mathbb{E}\left[\prod_{i=1}^n e^{tX_i}\right] = \prod_{i=1}^n \mathbb{E}\left[e^{tX_i}\right]$$

For any $t > 0$:

$$\begin{aligned}
\Pr[X > (1+\delta)\mu] &= \Pr[e^{tX} > e^{t(1+\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^n \mathbb{E}\left[e^{tX_i}\right]}{e^{t(1+\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^n (1 + p_i(e^t - 1))}{e^{t(1+\delta)\mu}} && \text{(By def of } \mathbb{E}\text{)} \\
&\leq \frac{\prod_{i=1}^n e^{p_i(e^t-1)}}{e^{t(1+\delta)\mu}} && \text{(By } 1 + x \leq e^x\text{)} \\
&= \frac{e^{\left(\sum_{i=1}^n p_i(e^t-1)\right)}}{e^{t(1+\delta)\mu}} \\
&= \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} = \left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^{\mu}
\end{aligned}$$

First Chernoff Bound, Proof

$$\mathbb{E}[e^{tX_i}] = p_i e^t + (1-p_i)e^0 = p_i e^t + 1 - p_i = 1 + p_i(e^t - 1)$$

For any $t > 0$:

$$\begin{aligned}
\Pr[X > (1+\delta)\mu] &= \Pr[e^{tX} > e^{t(1+\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^{n} \mathbb{E}\left[e^{tX_i}\right]}{e^{t(1+\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^{n}(1 + p_i(e^t - 1))}{e^{t(1+\delta)\mu}} && \text{(By def of } \mathbb{E}\text{)} \\
&\le \frac{\prod_{i=1}^{n} e^{p_i(e^t-1)}}{e^{t(1+\delta)\mu}} && \text{(By } 1 + x \le e^x\text{)} \\
&= \frac{e^{\left(\sum_{i=1}^{n} p_i(e^t-1)\right)}}{e^{t(1+\delta)\mu}} \\
&= \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} = \left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^{\mu}
\end{aligned}$$

First Chernoff Bound, Proof

For any $t > 0$:

$$\begin{aligned}
\Pr[X > (1+\delta)\mu] &= \Pr[e^{tX} > e^{t(1+\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^{n} \mathbb{E}\left[e^{tX_i}\right]}{e^{t(1+\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^{n}(1 + p_i(e^t - 1))}{e^{t(1+\delta)\mu}} && \text{(By def of } \mathbb{E}\text{)} \\
&\leq \frac{\prod_{i=1}^{n} e^{p_i(e^t-1)}}{e^{t(1+\delta)\mu}} && \text{(By } 1 + x \leq e^x\text{)} \\
&= \frac{e^{\left(\sum_{i=1}^{n} p_i(e^t-1)\right)}}{e^{t(1+\delta)\mu}} \\
&= \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} = \left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^{\mu}
\end{aligned}$$

First Chernoff Bound, Proof

For any $t > 0$:

$$
\begin{aligned}
\Pr[X > (1+\delta)\mu] &= \Pr[e^{tX} > e^{t(1+\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^{n} \mathbb{E}\left[e^{tX_i}\right]}{e^{t(1+\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^{n}(1 + p_i(e^t - 1))}{e^{t(1+\delta)\mu}} && \text{(By def of } \mathbb{E}\text{)} \\
&\leq \frac{\prod_{i=1}^{n} e^{p_i(e^t-1)}}{e^{t(1+\delta)\mu}} && \text{(By } 1 + x \leq e^x\text{)} \\
&= \frac{e^{\left(\sum_{i=1}^{n} p_i(e^t-1)\right)}}{e^{t(1+\delta)\mu}} \\
&= \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} = \left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^{\mu}
\end{aligned}
$$

First Chernoff Bound, Proof

For any $t > 0$:

$$\begin{aligned}
\Pr[X > (1+\delta)\mu] &= \Pr[e^{tX} > e^{t(1+\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^{n} \mathbb{E}\left[e^{tX_i}\right]}{e^{t(1+\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^{n}(1 + p_i(e^t - 1))}{e^{t(1+\delta)\mu}} && \text{(By def of } \mathbb{E}\text{)} \\
&\leq \frac{\prod_{i=1}^{n} e^{p_i(e^t-1)}}{e^{t(1+\delta)\mu}} && \text{(By } 1 + x \leq e^x\text{)} \\
&= \frac{e^{\left(\sum_{i=1}^{n} p_i(e^t-1)\right)}}{e^{t(1+\delta)\mu}} \\
&= \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} = \left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^{\mu}
\end{aligned}$$

First Chernoff Bound, Proof

For any $t > 0$:

$$\begin{aligned}
\Pr[X > (1+\delta)\mu] &= \Pr[e^{tX} > e^{t(1+\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^{n} \mathbb{E}\left[e^{tX_i}\right]}{e^{t(1+\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^{n}(1 + p_i(e^t - 1))}{e^{t(1+\delta)\mu}} && \text{(By def of } \mathbb{E}\text{)} \\
&\leq \frac{\prod_{i=1}^{n} e^{p_i(e^t-1)}}{e^{t(1+\delta)\mu}} && \text{(By } 1 + x \leq e^x\text{)} \\
&= \frac{e^{\left(\sum_{i=1}^{n} p_i(e^t-1)\right)}}{e^{t(1+\delta)\mu}} \\
&= \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} = \left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^{\mu}
\end{aligned}$$

First Chernoff Bound, Proof

We have shown that for any $t > 0$,

$$\Pr[X > (1+\delta)\mu] < \left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^{\mu}$$

Now find t minimizing this, by solving

$$\frac{d}{dt}\left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^{\mu} = 0$$

The solution happens to be $t = \ln(1+\delta)$, so

$$\Pr[X > (1+\delta)\mu] < \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu} \qquad \square$$

First Chernoff Bound, Proof

We have shown that for any $t > 0$,

$$\Pr[X > (1+\delta)\mu] < \left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^{\mu}$$

Now find t minimizing this, by solving

$$\frac{d}{dt}\left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^{\mu} = 0$$

The solution happens to be $t = \ln(1+\delta)$, so

$$\Pr[X > (1+\delta)\mu] < \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu} \qquad \square$$

First Chernoff Bound, Proof

We have shown that for any $t > 0$,

$$\Pr[X > (1+\delta)\mu] < \left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^\mu$$

Now find t minimizing this, by solving

$$\frac{d}{dt}\left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^\mu = 0$$

The solution happens to be $t = \ln(1+\delta)$, so

$$\Pr[X > (1+\delta)\mu] < \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\mu \qquad \square$$

$\frac{d}{dt}e^{\mu f(t)} = \mu f'(t)e^{\mu f(t)} = 0 \iff f'(t) = 0$, so we only need to solve
$\frac{d}{dt}((e^t - 1) - t(1+\delta)) = e^t - (1+\delta) = 0.$

First Chernoff Bound, Proof

We have shown that for any $t > 0$,

$$\Pr[X > (1+\delta)\mu] < \left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^{\mu}$$

Now find t minimizing this, by solving

$$\frac{d}{dt}\left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^{\mu} = 0$$

The solution happens to be $t = \ln(1+\delta)$, so

$$\Pr[X > (1+\delta)\mu] < \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu} \qquad \square$$

First Chernoff Bound, Example

Suppose you play a certain game n times, and win each time independently with probability $\frac{1}{3}$. What is the probability that you win more than half your games?

First Chernoff Bound, Example

Let $X = \sum_{i=1}^{n} X_i$, $\mu = \mathbb{E}[X] = \frac{n}{3}$, $\delta = \frac{1}{2}$.

$$\begin{aligned}
\Pr\left[X > \tfrac{n}{2}\right] &= \Pr[X > (1+\delta)\mu] \\
&< \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu} \\
&= \left(\frac{e^{\frac{1}{2}}}{(1+\frac{1}{2})^{(1+\frac{1}{2})}}\right)^{\frac{n}{3}} < (0.9646)^n
\end{aligned}$$

Note that this is *exponentially* small in n.

Contrast to Chebyshev, where

$\sigma_X^2 = \sum_{i=i}^{n} \sigma_{X_i}^2 = \sum_{i=1}^{n} p_i(1-p_i) \le \frac{n}{4}$, so $\sigma_X \le \frac{\sqrt{n}}{2}$

and $\Pr[X > \frac{n}{2}] \le \Pr[|X - \frac{n}{3}| \ge \frac{n}{6}] \le \frac{1}{\left(\frac{\sqrt{n}}{3}\right)^2} = \frac{9}{n}$.

First Chernoff Bound, Example

Let $X = \sum_{i=1}^{n} X_i$, $\mu = \mathbb{E}[X] = \frac{n}{3}$, $\delta = \frac{1}{2}$.

$$\begin{aligned}
\Pr\left[X > \tfrac{n}{2}\right] &= \Pr[X > (1+\delta)\mu] \\
&< \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu} \\
&= \left(\frac{e^{\frac{1}{2}}}{(1+\frac{1}{2})^{(1+\frac{1}{2})}}\right)^{\frac{n}{3}} < (0.9646)^n
\end{aligned}$$

Note that this is *exponentially* small in n.
Contrast to Chebyshev, where
$\sigma_X^2 = \sum_{i=i}^{n} \sigma_{X_i}^2 = \sum_{i=1}^{n} p_i(1-p_i) \le \frac{n}{4}$, so $\sigma_X \le \frac{\sqrt{n}}{2}$
and $\Pr[X > \frac{n}{2}] \le \Pr[|X - \frac{n}{3}| \ge \frac{n}{6}] \le \frac{1}{\left(\frac{\sqrt{n}}{3}\right)^2} = \frac{9}{n}$.

$$\frac{n}{2} - \frac{n}{3} = \frac{n}{6} = \frac{\sqrt{n}}{3}\frac{\sqrt{n}}{2} = \frac{\sqrt{n}}{3}\sigma_X.$$

Second Chernoff Bound

Theorem

Let $X_1, \ldots, X_n$ be independent Poisson trials such that, for $1 \leq i \leq n$, $\Pr[X_i = 1] = p_i$, where $0 < p_i < 1$. Then, for $X = \sum_{i=1}^{n} X_i$, $\mu = \mathbb{E}[X] = \sum_{i=1}^{n} p_i$, and any $0 < \delta < 1$.

$$\Pr[X < (1-\delta)\mu] < \left(\frac{e^{\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu}$$

Second Chernoff Bound

Theorem

Let $X_1, \ldots, X_n$ be independent Poisson trials such that, for $1 \leq i \leq n$, $\Pr[X_i = 1] = p_i$, where $0 < p_i < 1$. Then, for $X = \sum_{i=1}^{n} X_i$, $\mu = \mathbb{E}[X] = \sum_{i=1}^{n} p_i$, and any $0 < \delta < 1$.

$$\Pr[X < (1-\delta)\mu] < \left(\frac{e^{\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu}$$

$$< e^{-\frac{\delta^2 \mu}{2}}$$

Second Chernoff Bound, Proof

For any $t > 0$:

$$
\begin{aligned}
\Pr[X < (1-\delta)\mu] &= \Pr[e^{-tX} > e^{-t(1-\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{-tX}]}{e^{-t(1-\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^{n} \mathbb{E}\left[e^{-tX_i}\right]}{e^{-t(1-\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^{n}(1 + p_i(e^{-t} - 1))}{e^{-t(1-\delta)\mu}} && \text{(By def of } \mathbb{E}\text{)} \\
&\le \frac{\prod_{i=1}^{n} e^{p_i(e^{-t}-1)}}{e^{-t(1-\delta)\mu}} && \text{(By } 1 + x \le e^x\text{)} \\
&= \frac{e^{\left(\sum_{i=1}^{n} p_i(e^{-t}-1)\right)}}{e^{-t(1-\delta)\mu}} \\
&= \frac{e^{(e^{-t}-1)\mu}}{e^{-t(1-\delta)\mu}} = \left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu}
\end{aligned}
$$

Second Chernoff Bound, Proof

For any $t > 0$:

$$
\begin{aligned}
\Pr[X < (1-\delta)\mu] &= \Pr[e^{-tX} > e^{-t(1-\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{-tX}]}{e^{-t(1-\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^{n} \mathbb{E}\left[e^{-tX_i}\right]}{e^{-t(1-\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^{n} (1 + p_i(e^{-t} - 1))}{e^{-t(1-\delta)\mu}} && \text{(By def of } \mathbb{E}\text{)} \\
&\leq \frac{\prod_{i=1}^{n} e^{p_i(e^{-t}-1)}}{e^{-t(1-\delta)\mu}} && \text{(By } 1 + x \leq e^x\text{)} \\
&= \frac{e^{\left(\sum_{i=1}^{n} p_i(e^{-t}-1)\right)}}{e^{-t(1-\delta)\mu}} \\
&= \frac{e^{(e^{-t}-1)\mu}}{e^{-t(1-\delta)\mu}} = \left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu}
\end{aligned}
$$

Strict, because $0 < p_i < 1$ is strict.

Second Chernoff Bound, Proof

$$\mathbb{E}[e^{-tX}] = \mathbb{E}[e^{-t\sum_{i=1}^n X_i}] = \mathbb{E}\left[\prod_{i=1}^n e^{-tX_i}\right] = \prod_{i=1}^n \mathbb{E}\left[e^{-tX_i}\right]$$

For any $t > 0$:

$$\begin{aligned}
\Pr[X < (1-\delta)\mu] &= \Pr[e^{-tX} > e^{-t(1-\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{-tX}]}{e^{-t(1-\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^n \mathbb{E}\left[e^{-tX_i}\right]}{e^{-t(1-\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^n (1 + p_i(e^{-t} - 1))}{e^{-t(1-\delta)\mu}} && \text{(By def of } \mathbb{E}\text{)} \\
&\le \frac{\prod_{i=1}^n e^{p_i(e^{-t}-1)}}{e^{-t(1-\delta)\mu}} && \text{(By } 1 + x \le e^x\text{)} \\
&= \frac{e^{\left(\sum_{i=1}^n p_i(e^{-t}-1)\right)}}{e^{-t(1-\delta)\mu}} \\
&= \frac{e^{(e^{-t}-1)\mu}}{e^{-t(1-\delta)\mu}} = \left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu}
\end{aligned}$$

Second Chernoff Bound, Proof

$$\mathbb{E}[e^{-tX_i}] = p_i e^{-t} + (1-p_i)e^0 = p_i e^{-t} + 1 - p_i = 1 + p_i(e^{-t} - 1)$$

For any $t > 0$:

$$\begin{aligned}
\Pr[X < (1-\delta)\mu] &= \Pr[e^{-tX} > e^{-t(1-\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{-tX}]}{e^{-t(1-\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^{n} \mathbb{E}\left[e^{-tX_i}\right]}{e^{-t(1-\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^{n} (1 + p_i(e^{-t} - 1))}{e^{-t(1-\delta)\mu}} && \text{(By def of } \mathbb{E}\text{)} \\
&\le \frac{\prod_{i=1}^{n} e^{p_i(e^{-t}-1)}}{e^{-t(1-\delta)\mu}} && \text{(By } 1 + x \le e^x\text{)} \\
&= \frac{e^{\left(\sum_{i=1}^{n} p_i(e^{-t}-1)\right)}}{e^{-t(1-\delta)\mu}} \\
&= \frac{e^{(e^{-t}-1)\mu}}{e^{-t(1-\delta)\mu}} = \left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu}
\end{aligned}$$

Second Chernoff Bound, Proof

For any $t > 0$:

$$\begin{aligned}
\Pr[X < (1-\delta)\mu] &= \Pr[e^{-tX} > e^{-t(1-\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{-tX}]}{e^{-t(1-\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^{n} \mathbb{E}\left[e^{-tX_i}\right]}{e^{-t(1-\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^{n}(1 + p_i(e^{-t} - 1))}{e^{-t(1-\delta)\mu}} && \text{(By def of } \mathbb{E}\text{)} \\
&\le \frac{\prod_{i=1}^{n} e^{p_i(e^{-t}-1)}}{e^{-t(1-\delta)\mu}} && \text{(By } 1 + x \le e^x\text{)} \\
&= \frac{e^{\left(\sum_{i=1}^{n} p_i(e^{-t}-1)\right)}}{e^{-t(1-\delta)\mu}} \\
&= \frac{e^{(e^{-t}-1)\mu}}{e^{-t(1-\delta)\mu}} = \left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu}
\end{aligned}$$

Second Chernoff Bound, Proof

For any $t > 0$:

$$
\begin{aligned}
\Pr[X < (1-\delta)\mu] &= \Pr[e^{-tX} > e^{-t(1-\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{-tX}]}{e^{-t(1-\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^{n} \mathbb{E}\left[e^{-tX_i}\right]}{e^{-t(1-\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^{n}(1 + p_i(e^{-t} - 1))}{e^{-t(1-\delta)\mu}} && \text{(By def of } \mathbb{E}) \\
&\leq \frac{\prod_{i=1}^{n} e^{p_i(e^{-t}-1)}}{e^{-t(1-\delta)\mu}} && \text{(By } 1 + x \leq e^x) \\
&= \frac{e^{\left(\sum_{i=1}^{n} p_i(e^{-t}-1)\right)}}{e^{-t(1-\delta)\mu}} \\
&= \frac{e^{(e^{-t}-1)\mu}}{e^{-t(1-\delta)\mu}} = \left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu}
\end{aligned}
$$

Second Chernoff Bound, Proof

For any $t > 0$:

$$
\begin{aligned}
\Pr[X < (1-\delta)\mu] &= \Pr[e^{-tX} > e^{-t(1-\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{-tX}]}{e^{-t(1-\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^{n} \mathbb{E}\left[e^{-tX_i}\right]}{e^{-t(1-\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^{n}(1 + p_i(e^{-t} - 1))}{e^{-t(1-\delta)\mu}} && \text{(By def of } \mathbb{E}\text{)} \\
&\le \frac{\prod_{i=1}^{n} e^{p_i(e^{-t}-1)}}{e^{-t(1-\delta)\mu}} && \text{(By } 1 + x \le e^x\text{)} \\
&= \frac{e^{\left(\sum_{i=1}^{n} p_i(e^{-t}-1)\right)}}{e^{-t(1-\delta)\mu}} \\
&= \frac{e^{(e^{-t}-1)\mu}}{e^{-t(1-\delta)\mu}} = \left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu}
\end{aligned}
$$

Second Chernoff Bound, Proof

For any $t > 0$:

$$
\begin{aligned}
\Pr[X < (1-\delta)\mu] &= \Pr[e^{-tX} > e^{-t(1-\delta)\mu}] \\
&< \frac{\mathbb{E}[e^{-tX}]}{e^{-t(1-\delta)\mu}} && \text{(By Markov)} \\
&= \frac{\prod_{i=1}^{n} \mathbb{E}\left[e^{-tX_i}\right]}{e^{-t(1-\delta)\mu}} && \text{(By Independence)} \\
&= \frac{\prod_{i=1}^{n}(1 + p_i(e^{-t} - 1))}{e^{-t(1-\delta)\mu}} && \text{(By def of } \mathbb{E}\text{)} \\
&\le \frac{\prod_{i=1}^{n} e^{p_i(e^{-t}-1)}}{e^{-t(1-\delta)\mu}} && \text{(By } 1 + x \le e^x\text{)} \\
&= \frac{e^{\left(\sum_{i=1}^{n} p_i(e^{-t}-1)\right)}}{e^{-t(1-\delta)\mu}} \\
&= \frac{e^{(e^{-t}-1)\mu}}{e^{-t(1-\delta)\mu}} = \left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu}
\end{aligned}
$$

Second Chernoff Bound, Proof

We have shown that for any $t > 0$,

$$\Pr[X < (1-\delta)\mu] < \left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu}$$

Now find t minimizing this, by solving

$$\frac{d}{dt}\left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu} = 0$$

The solution happens to be $t = -\ln(1-\delta)$, so

$$\Pr[X < (1-\delta)\mu] < \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu}$$

Second Chernoff Bound, Proof

We have shown that for any $t > 0$,

$$\Pr[X < (1-\delta)\mu] < \left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu}$$

Now find t minimizing this, by solving

$$\frac{d}{dt}\left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu} = 0$$

The solution happens to be $t = -\ln(1-\delta)$, so

$$\Pr[X < (1-\delta)\mu] < \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu}$$

Second Chernoff Bound, Proof

We have shown that for any $t > 0$,

$$\Pr[X < (1-\delta)\mu] < \left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu}$$

Now find t minimizing this, by solving

$$\frac{d}{dt}\left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu} = 0$$

The solution happens to be $t = -\ln(1-\delta)$, so

$$\Pr[X < (1-\delta)\mu] < \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu}$$

$\frac{d}{dt}e^{\mu f(t)} = \mu f'(t)e^{\mu f(t)} = 0 \iff f'(t) = 0$, so we only need to solve $\frac{d}{dt}((e^{-t}-1) + t(1-\delta)) = -e^{-t} + (1-\delta) = 0$.

So $t = -\ln(1-\delta)$.

Second Chernoff Bound, Proof

We have shown that for any $t > 0$,

$$\Pr[X < (1-\delta)\mu] < \left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu}$$

Now find t minimizing this, by solving

$$\frac{d}{dt}\left(\frac{e^{(e^{-t}-1)}}{e^{-t(1-\delta)}}\right)^{\mu} = 0$$

The solution happens to be $t = -\ln(1-\delta)$, so

$$\Pr[X < (1-\delta)\mu] < \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu}$$

This bound is completely analoguous to the first Chernoff bound. However, it turns out there is a slightly weaker bound that is often more useful.

Second Chernoff Bound, Proof

Now use that for $-1 < x < 1$

$$
\begin{aligned}
\ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} + \cdots \qquad \text{(McLaurin)} \\
\ln(1-x) &= (-x) - \frac{(-x)^2}{2} + \frac{(-x)^3}{3} + \cdots \\
-\ln(1-x) &= x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots > x + \frac{x^2}{2}
\end{aligned}
$$

so

$$
\begin{aligned}
(1-\delta)^{(1-\delta)} &= e^{(1-\delta)\cdot\ln(1-\delta)} \\
&= e^{(-(1-\delta))\cdot(-\ln(1-\delta))} \\
&> e^{(\delta-1)\cdot(\delta+\frac{\delta^2}{2})} \\
&= e^{-\delta+\frac{\delta^2}{2}+\frac{\delta^3}{2}} \\
&> e^{-\delta+\frac{\delta^2}{2}}
\end{aligned}
$$

Second Chernoff Bound, Proof

We have shown that

$$
\begin{aligned}
\Pr[X < (1-\delta)\mu] &< \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu} \\
&< \left(\frac{e^{-\delta}}{e^{-\delta+\frac{\delta^2}{2}}}\right)^{\mu} \\
&= e^{-\frac{\delta^2\mu}{2}} \qquad \square
\end{aligned}
$$

Second Chernoff Bound, Proof

We have shown that

$$
\begin{aligned}
\Pr[X < (1-\delta)\mu] &< \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu} \\
&< \left(\frac{e^{-\delta}}{e^{-\delta+\frac{\delta^2}{2}}}\right)^{\mu} \\
&= e^{-\frac{\delta^2 \mu}{2}} \qquad \square
\end{aligned}
$$

Second Chernoff Bound, Proof

We have shown that

$$\begin{aligned}
\Pr[X < (1-\delta)\mu] &< \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu} \\
&< \left(\frac{e^{-\delta}}{e^{-\delta+\frac{\delta^2}{2}}}\right)^{\mu} \\
&= e^{-\frac{\delta^2\mu}{2}} \qquad \square
\end{aligned}$$

Second Chernoff Bound, Example

Suppose you play a certain game n times, and win each time independently with probability $\frac{3}{4}$. What is the probability that you win less than half your games?

First Chernoff Bound, Example

Let $X = \sum_{i=1}^{n} X_i$, $\mu = \mathbb{E}[X] = \frac{3n}{4}$, $\delta = \frac{1}{3}$.

$$\begin{aligned}
\Pr\left[X < \frac{n}{2}\right] &= \Pr[X < (1-\delta)\mu] \\
&< e^{-\frac{\delta^2 \mu}{2}} \\
&= e^{-\frac{\frac{3n}{4}\left(\frac{1}{3}\right)^2}{2}} = e^{-\frac{1}{24}n} < (0.9592)^n
\end{aligned}$$

Note again that this is *exponentially* small in n.

Notation and Computing Deviations

Define

$$F^{-}(\mu, \delta) := e^{-\frac{\delta^2 \mu}{2}}$$
$$\Delta^{-}(\mu, \epsilon) := \min \delta \text{ such that } F^{-}(\mu, \delta) \leq \epsilon$$

From the definition it is easy to compute

$$\Delta^{-}(\mu, \epsilon) = \sqrt{\frac{2 \ln(\frac{1}{\epsilon})}{\mu}}$$

Notation and Computing Deviations

Similarly, define

$$F^+(\mu,\delta) := \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\mu$$

$$\Delta^+(\mu,\epsilon) := \min \delta \text{ such that } F^+(\mu,\delta) \le \epsilon$$

This is not easy to calculate. One attempt is

$$\Delta^+(\mu,\epsilon) = \min \begin{cases} \delta_1 = \sqrt{\frac{4\ln\frac{1}{\epsilon}}{\mu}} & \text{if } 0 < \delta_1 \le 2e-1 \\ \delta_2 = \frac{\log_2\frac{1}{\epsilon}}{\mu} - 1 & \text{if } \delta_2 > 2e-1 \end{cases}$$

The funky notation here means that each of δ_1, δ_2 is only a valid candidate if it falls in the specified range.

We will not prove that this works.

Notation and Computing Deviations

Similarly, define

$$F^+(\mu,\delta) := \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\mu \le e^{-\frac{\delta^2\mu}{2+\delta}}$$

$$\Delta^+(\mu,\epsilon) := \min \delta \text{ such that } F^+(\mu,\delta) \le \epsilon$$

Using the simpler upper bound we can find

$$\Delta^+(\mu,\epsilon) \le \frac{1}{2\mu}\left(\ln\frac{1}{\epsilon} + \sqrt{8\mu + \ln\frac{1}{\epsilon}} \cdot \sqrt{\ln\frac{1}{\epsilon}}\right)$$

Wikipedia lists several bounds on F^+ that can help, for example this, based on the inequality $\frac{2\delta}{2+\delta} \le \ln(1+\delta)$.

Summary

- We Analyzed the Coupon Collectors Problem.
- We Proved two Chernoff Bounds.
- Tomorrow: Two applications of Chernoff Bounds.

Summary

- We Analyzed the Coupon Collectors Problem.
- We Proved two Chernoff Bounds.
- Tomorrow: Two applications of Chernoff Bounds.

Summary

- We Analyzed the Coupon Collectors Problem.
- We Proved two Chernoff Bounds.
- Tomorrow: Two applications of Chernoff Bounds.

Summary

- We Analyzed the Coupon Collectors Problem.
- We Proved two Chernoff Bounds.
- Tomorrow: Two applications of Chernoff Bounds.