Randomness in Computing

LECTURE 13

CS 537

Last time

- Finished routing on hypercube
- Balls-and-Bins model
  - Birthday Paradox
  - Application: Bucket Sort

Today

- Poisson distribution
- Poisson approximation

Balls and Bins in the Limit

m balls into n bins

- The probability that bin 1 is empty is

$$\text{for } |x| \le 1$$

$$e^{-x}(1-x^2) \le 1-x \le e^{-x}$$

- The probability p_k that bin 1 has k balls is

$$p_k = \binom{m}{k}\left(\frac{1}{n}\right)^k\left(1-\frac{1}{n}\right)^{m-k} = \frac{m!}{k!\,(m-k)!}\frac{1}{n^k}\left(1-\frac{1}{n}\right)^{m-k}$$

$$= \frac{1}{k!}\left(\frac{m}{n}\cdot\frac{m-1}{n}\cdots\frac{m-k+1}{n}\right)\left(1-\frac{1}{n}\right)^{m-k}$$

Poisson random variables

- A Poisson random variable with parameter μ is given by the following distribution on $j = 0,1,2,\ldots$

$$\Pr[X = j] = \frac{e^{-\mu}\mu^j}{j!}$$

- Check that probabilities sum to 1:

Taylor expansion: $e^x = \sum_{j=0}^{\infty} \frac{x^j}{j!}$

$$\sum_{j=0}^{\infty} \Pr[X = j] = \sum_{j=0}^{\infty} \frac{e^{-\mu}\mu^j}{j!} =$$

- The expectation of a Poisson R.V. X is

$\mathbb{E}[X] =$

$\mathrm{var}[X] = \mu$ (See Ex. 5.5)

Independent Poisson RVs

Theorem

Let X and Y be independent Poisson RVs with means μ_X and μ_Y.

Then $X+Y$ is a Poisson RV with mean $\mu_X+\mu_Y$.

Chernoff Bounds for Poisson RVs

Theorem. Let X be a Poisson RV with mean μ.

- (upper tail, additive) If $x > \mu$, then
$$\Pr[X \geq x] \leq \frac{e^{-\mu}(e\mu)^x}{x^x}.$$
- (lower tail, additive) If $x < \mu$, then
$$\Pr[X \leq x] \leq \frac{e^{-\mu}(e\mu)^x}{x^x}.$$
- (upper tail, multiplicative) For any $\delta > 0$,
$$\Pr[X \geq (1+\delta)\mu] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}.$$
- (lower tail, multiplicative) For any $\delta \in (0,1)$,
$$\Pr[X \leq (1-\delta)\mu] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}.$$

Poisson Distribution is Limit of Binomial Distribution

Theorem

Let $X_n \sim \text{Bin}(n, p)$, where p is a function of n and $\lim_{n\to\infty} np = \mu$, a constant independent of n.

Then, for all fixed k,

$$\lim_{n\to\infty} \Pr[X_n = k] = \frac{e^{-\mu}\mu^k}{k!}.$$

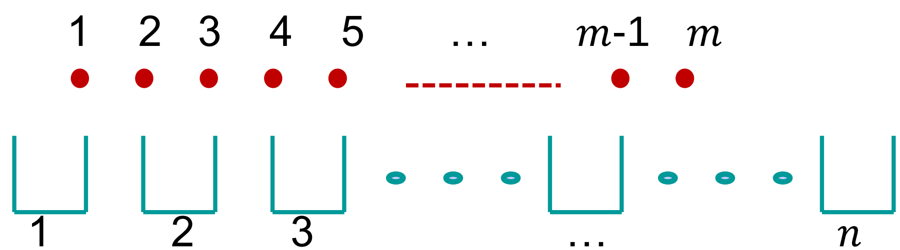
- Applies to balls-and-bins if $m = nc$.

The Poisson Approximation

- The Balls-and-Bins model has dependences.
- E.g. if Bin 1 is empty, then Bin 2 is less likely to be empty.
- The Poisson Approximation gets rid of dependencies.

Poisson approximation

- The Balls-and-Bins model has dependences.
- The Poisson approximation gets rid of dependencies.



- m balls into n bins u.i.r.

  For $i \in [n]$, let

  (real world) $X_i^{(m)} =$ # of balls in bin i

  (Poisson world) $Y_i^{(m)} \sim Poisson(\mu)$, where $\mu = \frac{m}{n}$ and

  $Y_i^{(m)}$ are mutually independent.

- If we condition the Poisson distribution on producing exactly k balls, then it's the same as the distribution resulting from throwing k balls into n bins.

Poisson Distribution Conditioned on getting k balls

Poisson Distribution Theorem

Poisson world

The distribution of $\left(Y_1^{(m)}, \dots, Y_n^{(m)}\right)$ conditioned on $\sum_{i\in[n]} Y_i^{(m)} = k$ is the same as $\left(X_1^{(k)}, \dots, X_n^{(k)}\right)$, regardless of the value of m.

Real world

Proof: Consider any $k_1, \dots, k_n$ satisfying $\sum_{i\in[n]} k_i = k$

- $\Pr\left[\left(X_1^{(k)}, \dots, X_n^{(k)}\right) = (k_1, \dots, k_n)\right]$ is

Multinomial coefficient

$$\frac{\binom{k}{k_1, \dots, k_n}}{n^k} = \frac{k!}{k_1!\, k_2! \cdots k_n!} \cdot \frac{1}{n^k}$$

Poisson Distribution Conditioned on getting k balls

Poisson Distribution Theorem

Poisson world

The distribution of $\left(Y_1^{(m)}, \dots, Y_n^{(m)}\right)$ conditioned on $\sum_{i\in[n]} Y_i^{(m)} = k$ is the same as $\left(X_1^{(k)}, \dots, X_n^{(k)}\right)$, regardless of the value of m.

Real world

Proof: Consider any $k_1, \dots, k_n$ satisfying $\sum_{i\in[n]} k_i = k$

- $\Pr\left[\left(X_1^{(k)}, \dots, X_n^{(k)}\right) = (k_1, \dots, k_n)\right]$ is $\dfrac{k!}{k_1!\,k_2!\cdots k_n!} \cdot \dfrac{1}{n^k}$
- $\Pr\left[\left(Y_1^{(m)}, \dots, Y_n^{(m)}\right) = (k_1, \dots, k_n) \,\middle|\, \sum_{i\in[n]} Y_i^{(m)} = k\right]$ is

$$\frac{\Pr\left[\left(Y_1^{(m)} = k_1\right) \cap \dots \cap \left(Y_n^{(m)} = k_n\right)\right]}{\Pr\left[\sum_{i\in[n]} Y_i^{(m)} = k\right]} = \frac{\prod_{i\in[n]}}{\dfrac{e^{-m}\, m^k}{k!}}$$

Poisson RV

Approximating a function of the loads of the bins

Poisson Approximation Theorem

Let $f(x_1, \dots, x_n) \geq 0$ for all $x_1, \dots, x_n \in \{0,1,2,\dots\}$. Then

$$\mathbb{E}\left[f\left(X_1^{(m)}, \dots, X_n^{(m)}\right)\right] \leq \boldsymbol{e\sqrt{m}} \cdot \mathbb{E}\left[f\left(Y_1^{(m)}, \dots, Y_n^{(m)}\right)\right].$$

exact case ***Poisson case***

- Fact ($Stirling's\ formula$): $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$
- Bounds for all $n \in \mathbb{N}$: $\sqrt{2\pi}\sqrt{n}\left(\frac{n}{e}\right)^n \leq n! \leq e\sqrt{n}\left(\frac{n}{e}\right)^n$

Approximating a function of the loads of the bins

Poisson Approximation Theorem

Let $f(x_1, \dots, x_n) \geq 0$ for all $x_1, \dots, x_n \in \{0,1,2, \dots\}$. Then

$$\mathbb{E}\left[f\left(X_1^{(m)}, \dots, X_n^{(m)}\right)\right] \leq e\sqrt{m} \cdot \mathbb{E}\left[f\left(Y_1^{(m)}, \dots, Y_n^{(m)}\right)\right].$$

Poisson case

Proof: $\mathbb{E}\left[f\left(Y_1^{(m)}, \dots, Y_n^{(m)}\right)\right]$

$$= \sum_{k=0}^{\infty} \mathbb{E}\left[f\left(Y_1^{(m)}, \dots, Y_n^{(m)}\right) \mid \sum_{i\in[n]} Y_i^{(m)} = k\right] \cdot \Pr\left[\sum_{i\in[n]} Y_i^{(m)} = k\right]$$

Law of Total Expectation

$$\geq \mathbb{E}\left[f\left(Y_1^{(m)}, \dots, Y_n^{(m)}\right) \mid \sum_{i\in[n]} Y_i^{(m)} = m\right] \cdot \Pr\left[\sum_{i\in[n]} Y_i^{(m)} = m\right]$$

$$= \mathbb{E}\left[f\left(X_1^{(m)}, \dots, X_n^{(m)}\right)\right] \cdot \Pr[Y = m]$$

Approximating a function of the loads of the bins

Poisson Approximation Theorem

Let $f(x_1, \ldots, x_n) \geq 0$ for all $x_1, \ldots, x_n \in \{0,1,2, \ldots\}$. Then

$$\mathbb{E}\left[f\left(X_1^{(m)}, \ldots, X_n^{(m)}\right)\right] \leq e\sqrt{m} \cdot \mathbb{E}\left[f\left(Y_1^{(m)}, \ldots, Y_n^{(m)}\right)\right].$$

- Poisson case: # of balls in each bin is independent Poisson$\left(\frac{m}{n}\right)$
- Corollary. Any event that has probability p in the Poisson case has probability $\leq p \cdot e\sqrt{m}$ in the exact case.

  Proof: Let X be the indicator for that event.

  Then $\mathbb{E}[X]$ is the probability that event occurs.
- Improvements to Theorem and Corollary

  If $\mathbb{E}\left[f\left(X_1^{(m)}, \ldots, X_n^{(m)}\right)\right]$ is monotonically nonincreasing (or nondecreasing) in m, then $e\sqrt{m}$ can be changed to 2.

Application: Max Load

n balls into n bins

sufficiently large

- Before (by Chernoff): $\Pr\left[MaxLoad > \frac{3\ln n}{\ln\ln n}\right] \le \frac{1}{n}$ for s.l. n
- Theorem. $\Pr\left[MaxLoad < \frac{\ln n}{\ln\ln n}\right] \le \frac{1}{n}$ for s.l. n

Proof: Let $M = \frac{\ln n}{\ln\ln n}$