CSTH - Spring 2024
Lecture 18 - March 19
Summary of Last Lecture - all based on balls & bins

1. For a random hash function, to avoid collisions (with good prob.), need size of hash table to be
   \[ \approx (\text{no. of keys stored})^2 \]

2. To collect at least one copy of each of \( n \) coupons, need to take about \( n \ln n \) random samples

3. If we randomly distribute \( n \) jobs among \( n \) processors, the largest load on any processor is likely to be around
   \[ \frac{\ln n}{\ln \ln n} \]
Today

- Random variables (= functions/measurements on probability spaces)
- Distributions
- Expectation
- The Unreasonable Power of Linearity of Expectation
Random Variables
Measurements on probability spaces

Example: \( \mathcal{R} = \) space of 3 fair coin tosses
\[
= \{ \text{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT} \}
\]

Uniform probabilities: \( \Pr[\omega] = \frac{1}{8} \) \( \forall \omega \in \mathcal{R} \)

For any \( \omega \in \mathcal{R} \), let \( X(\omega) := \) number of Heads in \( \omega \)

Note:
- \( X(\omega) \) is a (real) number (actually, a non-negative integer)
- \( X(\omega) \in \{0, 1, 2, 3\} \)
For any \( \omega \in \Omega \), let \( X(\omega) := \text{number of Heads in } \omega \)

- For any \( i \in \{0, 1, 2, 3\} \), \( E_i := \{ \omega : X(\omega) = i \} \) is an event

- The events \( \{ E_i \} \) partition \( \Omega \)

The collection \( \{(i, \Pr[E_i])\} \) is the distribution of \( X \)

Here:

\[
\begin{align*}
\Pr[X = 0] &= \frac{1}{8} \\
\Pr[X = 1] &= \frac{3}{8} \\
\Pr[X = 2] &= \frac{3}{8} \\
\Pr[X = 3] &= \frac{1}{8}
\end{align*}
\]
Random Variable: Definition

Defn: A random variable on a prob. space $\Omega$ is a function $X: \Omega \rightarrow \mathbb{R}$.

I.e., $X$ assigns a real value $X(w)$ to each $w \in \Omega$.

Defn: The distribution of a (discrete) r.v. $X$ is:
- the set of possible values for $X$
- the probability $\Pr[X = a]$ for each possible value $a$.
Check: For any r.v. $X$ with set of possible values $\mathcal{A}$, we have
\[
\sum_{a \in \mathcal{A}} \Pr[X = a] = 1
\]

Proof: $\sum_{a \in \mathcal{A}} \Pr[X = a] = \sum_{a \in \mathcal{A}} \left(\sum_{\omega : X(\omega) = a} \Pr[\omega]\right)$
\[
= \sum_{\omega \in \Omega} \Pr[\omega] = 1 \quad \Box
\]

since $\forall \omega \in \Omega$
\[\exists \text{ unique } a \in \mathcal{A} \text{ s.t. } X(\omega) = a\]
Examples

1. Roll 2 fair dice
   X = sum of scores on dice
   \( \Omega = \{(i,j): 1 \leq i, j \leq 6\} \quad |\Omega| = 36 \)
   \[ X(i,j) = i+j \]
   \( X(\omega) \in \{2, 3, \ldots, 11, 12\} \)

Distribution of X

Pr [X=2] = \( \frac{1}{36} \)
Pr [X=3] = \( \frac{2}{36} = \frac{1}{18} \)
Pr [X=4] = \( \frac{3}{36} = \frac{1}{12} \)
Pr [X=5] = \( \frac{4}{36} = \frac{1}{9} \)
Pr [X=6] = \( \frac{5}{36} \)
Pr [X=7] = \( \frac{6}{36} = \frac{1}{6} \)
Pr [X=8] = \( \frac{5}{36} \)
Pr [X=9] = \( \frac{4}{36} = \frac{1}{9} \)
Pr [X=10] = \( \frac{3}{36} = \frac{1}{12} \)
Pr [X=11] = \( \frac{2}{36} = \frac{1}{18} \)
Pr [X=12] = \( \frac{1}{36} \)
2. **Random Permutations**

Collect the IDs of $n$ students
Redistribute them randomly (one per student)

$\Omega = \text{set of permutations of } n \text{ items}$

$|\Omega| = n!$

E.g. for $n=3$

Uniform probability space: $P(\omega) = \frac{1}{n!}$ \forall $\omega$

**Random variable** $X = \text{no. of students who get their own ID}$

a.k.a. "fixed points"
$X = \text{no. of fixed points in a random permutation}$

Distribution of $X$

- $P_r[X = 0] = \frac{2}{6} = \frac{1}{3}$
- $P_r[X = 1] = \frac{3}{6} = \frac{1}{2}$
- $P_r[X = 3] = \frac{1}{6} = \frac{1}{6}$

Histogram

- $P_r[X = a]$
  - $\frac{1}{3}$ for $a = 0$
  - $\frac{1}{2}$ for $a = 1$
  - $\frac{1}{6}$ for $a = 3$
3. Binomial Distribution

Toss \( n \) biased coins, each having Heads probability \( p \).

\[ \Omega = \{ H, T \}^n \] (all strings of length \( n \) over alphabet \( \{ H, T \} \))

\[ \Pr[\omega] = p^i (1-p)^{n-i} \]

where \( i = \text{no. of Heads in } \omega \)

Random variable \( X = \text{no. of Heads} \), \( X \in \{0,1,\ldots,n\} \)

What is the distribution of \( X \)?

\[ \Pr[X=i] = \binom{n}{i} p^i (1-p)^{n-i} \]

We say \( X \) has binomial distribution with parameters \( n, p \):

\[ X \sim \text{Bin}(n, p) \]
Pictures of Bin \((n, p)\)

Binomial distribution
\(n = 9, \ p = \frac{1}{2}\)

Binomial distribution
\(n = \text{large}, \ p = \text{small}\)
4. **Hypergeometric Distribution**

Deal a 5-card poker hand \( \mathcal{H}_2 \) = \( \binom{52}{5} \)

R.v. \( X = \) no. of hearts in your hand

\( X \subseteq \{0, 1, 2, 3, 4, 5\} \)

**Distribution:**

\[
\begin{align*}
P_X[X=0] &= \frac{\binom{39}{5}}{\binom{52}{5}} \\
\end{align*}
\]

\[
\begin{align*}
P_X[X=1] &= \frac{(\binom{13}{1}) \cdot \binom{39}{4}}{\binom{52}{5}} \\
\end{align*}
\]

\[
\begin{align*}
&\vdots \\
P_X[X=k] &= \frac{(\binom{13}{k}) \cdot \binom{39}{5-k}}{\binom{52}{5}} \\
\end{align*}
\]
4. **Hypergeometric Distribution**

Deal a 5-card poker hand $\mathcal{D} = \binom{52}{5}$

R.v. $X =$ no. of hearts in your hand

$X \in \{0, 1, 2, 3, 4, 5\}$

**Distribution:**

\[
\begin{align*}
\Pr[X = 0] &= \frac{\binom{39}{5}}{\binom{52}{5}} \\
\Pr[X = 1] &= \frac{\binom{13}{1} \times \binom{39}{4}}{\binom{52}{5}} \\
\quad \vdots \\
\Pr[X = k] &= \frac{\binom{13}{k} \times \binom{39}{5-k}}{\binom{52}{5}} 
\end{align*}
\]

**Note:**

\[
\sum_{k=0}^{n} \binom{B}{k} \binom{N-B}{n-k} = \binom{N}{n}
\]

Move generally:

- box of $N$ balls, $B$ black, rest white
- draw $n$ balls w/o replacement
- $X =$ # of black balls drawn

\[
\Pr[X = k] = \frac{\binom{B}{k} \binom{N-B}{n-k}}{\binom{N}{n}}
\]

**Hypergeometric distribution**, parameters $(N, B, n)$
Joint Distributions

Defn: The joint distribution of two r.v.'s $X, Y$ on the same prob. space is the set
\[
\{(a, b, \Pr[X=a, Y=b] : a \in A, b \in B)\}
\]
where $A, B$ are the possible values of $X, Y$ resp.

The marginal distribution of $X$ is given by
\[
\Pr[X=a] = \sum_{b \in B} \Pr[X=a, Y=b]
\]

$X, Y$ are independent if
\[
\Pr[X=a, Y=b] = \Pr[X=a] \times \Pr[Y=b] \quad \forall a, b
\]
Joint Distributions

Example: Throw two fair dice

Random variables:

\[
X = \text{score on first die} \quad Y = \text{score on second} \quad Z = \text{sum of scores}
\]

\[
\Pr[X=3, Y=5] = \frac{1}{36}
\]

\[
\Pr[X=3, Z=9] = \frac{1}{36}
\]

Are \( X, Y \) independent? Are \( X, Z \) independent?
**Expectation**  (= mean/average)

Simplest quantity that summarizes the distribution of a r.v.

**Defn:** The expectation of a (discrete) r.v. $X$ is

$$E[X] := \sum_{a \in A} a \times \Pr[X = a]$$

where $A$ is the set of possible values of $X$

$E[X]$ measures the "center of mass" of the distribution
Expectation: Examples

1. $X = \text{score on one fair die}$

\[
E[X] = \left( \frac{1}{6} \times 1 \right) + \left( \frac{1}{6} \times 2 \right) + \ldots + \left( \frac{1}{6} \times 6 \right)
\]

\[
= \frac{1}{6} \times (1 + 2 + \ldots + 6) = \frac{7}{2}
\]

1'. $Y = \text{sum of scores on two fair dice}$

\[
E[Y] = \left( \frac{1}{36} \times 2 \right) + \left( \frac{2}{36} \times 3 \right) + \left( \frac{3}{36} \times 4 \right) + \ldots + \left( \frac{1}{36} \times 12 \right)
\]

\[
= \ldots = \boxed{7}
\]
2. **$X$ = no. of fixed points in a random permutation ($n=3$)**

**Distribution of $X$**

- $\Pr[X=0] = \frac{2}{6} = \frac{1}{3}$
- $\Pr[X=1] = \frac{3}{6} = \frac{1}{2}$
- $\Pr[X=3] = \frac{1}{6} = \frac{1}{6}$

\[
E[X] = \left( \frac{1}{3} \times 0 \right) + \left( \frac{1}{2} \times 1 \right) + \left( \frac{1}{6} \times 3 \right) = 0 + \frac{1}{2} + \frac{1}{2} = 1
\]
3. **Roulette**

Roulette wheel: 36 numbers (18 black/18 red) plus 0, 00

Bet $1 on red: win $1 if red, lose $1 if black or green

\[ X = \text{amount won lost} \]

\[
\begin{align*}
Pr(X = +1) &= \frac{18}{38} \\
Pr(X = -1) &= \frac{20}{38}
\end{align*}
\]

\[
E[X] = (1 \times \frac{18}{38}) + (-1 \times \frac{20}{38}) = -\frac{1}{19}
\]

**Note:** presence of 0, 00 make this game unfair
Linearity of Expectation

Thm: For any random variables $X, Y$ on prob. space $\mathcal{F}$,

(i) $E[X+Y] = E[X] + E[Y]$

(ii) $E[aX] = aE[X]$ for constant $a$

Proof: Note that $E[X] = \sum_{\omega \in \mathcal{F}} X(\omega) \times \Pr[\omega]$

(i) $E[X+Y] = \sum_{\omega \in \mathcal{F}} (X+Y)(\omega) \times \Pr[\omega]$

$= \sum_{\omega \in \mathcal{F}} X(\omega) \times \Pr[\omega] + \sum_{\omega \in \mathcal{F}} Y(\omega) \times \Pr[\omega]$

$= E[X] + E[Y]$

(ii) Easy exercise

Crucial: Does not assume $X, Y$ are independent!!!

Linearity of Expectation: Examples

1. Two fair dice
   \[ X = \text{sum of dice rolls} \]
   Then \[ X = X_1 + X_2 \] where \( X_1 = \text{score on first die} \)
   \[ X_2 = \text{second} \]

   For \( \omega \in \Omega \):
   \[ X(\omega) = X_1(\omega) + X_2(\omega) \]

   So by linearity
   \[ E[X] = E[X_1 + X_2] = E[X_1] + E[X_2] = \frac{7}{2} + \frac{7}{2} = 7 \]
2. **Multiple roulette games**

Play roulette 100 times (\$1 stake each time)

\[ X = \text{amount won/lost} \]

\[ X = X_1 + X_2 + \ldots + X_{100} \]

where \( X_i = \text{amt. won/lost in } i\text{th game} \)

Recall: \( E[X_i] = -\frac{1}{19} \quad \forall i \)

**Linearity:**

\[ E[X] = \sum_{i=1}^{100} E[X_i] = 100 \times \left(-\frac{1}{19}\right) \approx -5.26 \]

**Summary:** The expected value of the sum of 100 roulette games is approximately -5.26, indicating a slight disadvantage to the player over the long run.
3. **Multiple coin tosses**

Toss a biased coin (Heads prob. \( p \)) \( n \) times

\[
X = \# \text{Heads} \quad X \sim \text{Binomial}(n, p)
\]

\[
X = X_1 + \cdots + X_n \quad \text{where} \quad X_i = \begin{cases} 1 & \text{if ith toss Heads} \\ 0 & \text{Tails} \end{cases}
\]

Note that \( E[X_i] = (1 \times p) + (0 \times (1-p)) = p \)

**Linearity:** \( E[X] = \sum_{i=1}^{n} E[X_i] = n \times p \)
4. **Balls & Bins**

Recall: toss $m$ balls u.a.r. into $n$ bins

R.v. $X = \#$ of empty bins

\[ X = \sum_{i=1}^{n} X_i \]

where $X_i = \begin{cases} 1 & \text{if bin } i \text{ empty} \\ 0 & \text{if bin } i \text{ not empty} \end{cases}$

Then $E[X_i] = (1 \times \Pr[\text{bin } i \text{ empty}]) + (0 \times \Pr[\text{bin } i \text{ not empty}])$

\[ = \Pr[\text{bin } i \text{ empty}] \]

\[ = (1 - \frac{1}{n})^m \]

Linearity:

\[ E[X] = \sum_{i=1}^{n} E[X_i] = n (1 - \frac{1}{n})^m \approx ne^{-m/n} \]

E.g. if $m=n$, $E[X] \approx ne^{-1} \approx 0.37n$
5. **Fixed points in a random permutation**

**General case: n items**

\[ X_n = \# \text{ of fixed points} \]

\[ X = \sum_{i=1}^{n} X_i \quad \text{where} \quad X_i = \begin{cases} 1 & \text{if } i \text{ is a fixed point} \\ 0 & \text{otherwise} \end{cases} \]

\[ E[X_i] = \Pr [i \text{ is a fixed point}] = \frac{1}{n} \]

**Linearity:**

\[ E[X] = \sum_{i=1}^{n} E[X_i] = n \times \frac{1}{n} = 1 \]

**Bottom line:** If we collect & redistribute IDs of \( n \) people, the expected \# who get their own ID is **always 1** (indep. of \( n \))