Last time:

Markov’s ineq.: \[ Pr\left(X \geq a\right) \leq \frac{E(X)}{a} \] if \( X \geq 0 \)

Chebychev’s ineq.: \[ Pr\left(|X - E(X)| \geq a\right) \leq \frac{Var(X)}{a^2} \] (always)

Weak law of lg. #s: \[ Pr\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| \geq a\right) \leq \frac{\sigma^2}{na^2} \] if \( X_i \) indep. and \( E(X_i) = \mu, Var(X_i) = \sigma^2 \)

Announcements:

- TA applications open
Defn: a $T$-valued RV is a function from $S$ to $T$

For our examples today:

- We are putting names (strings) in buckets from 1 to $m$
- $X := \{Alice, Andrew, Jehron, JiHun, \ldots \}$
- $Y := \{1, \ldots, 26\}$ is the set of buckets (more generally $\{1, \ldots, m\}$)
- $S$ is the sample space.

Defn: a family of hash functions $H$ is a $[X \to Y]$-valued RV

- i.e. a function $H : S \to [X \to Y]$
- in the context of hashing, outcomes are referred to as seeds
- the outputs of $H$ (i.e. functions $h : X \to Y$) are called hash functions

$H$ gives a separate $Y$-valued RV $H_x$ for each $x \in X$

- $H_x : S \to Y$ is given by $H_x(s) := (H(s))(x)$

$H$ is nice if

1. For any $x$ and $y$, $Pr(H_x = y) = \frac{1}{|Y|}$.
2. For any $x_1 \neq x_2$, $H_{x_1}$ and $H_{x_2}$ are independent.