Outline

Balls and Bins (and Birthdays and Coupons!)

Clock Solitaire and Principle of Deferred Decisions
Balls and Bins

Throw $m$ balls into $n$ bins where each throw is independent. Many questions:

- The maximum number of balls that fall into the same bin? (Load Balancing)
- How large must $m$ be such that there exists a bin with at least two balls? (Birthday Paradox)
- How large must $m$ be such that all bins get at least one ball? (Coupon Collecting)
Heaviest Bin (1/2)

Assume $m = n$. Let $Y_i$ be number of balls that fall in $i$-th bin.

**Lemma**

Let $k \geq (3 \ln n)/\ln \ln n$. Then $\Pr[Y_i \geq k] \leq n^{-2}$

**Proof.**

- $\Pr[Y_i = j] = \binom{n}{j}(1/n)^j(1 - 1/n)^{n-j}$
- Using the bound $\binom{n}{j} \leq (ne/j)^j$:
  $$\Pr[Y_i = j] = \binom{n}{j}(1/n)^j(1 - 1/n)^{n-j} \leq (e/j)^j$$
- By summing up a geometric series:
  $$\Pr[Y_i \geq k] = \sum_{j \geq k} (e/j)^j \leq (e/k)^k \frac{1}{1 - e/k}$$

\[\square\]
Assume $m = n$. Let $Y_i$ be number of balls that fall in $i$-th bin.

**Lemma**

Let $k \geq (3 \ln n) / \ln \ln n$. Then $\mathbb{P} [Y_i \geq k] \leq n^{-2}$

**Theorem**

$\mathbb{P} [Y_i < k$ for all $i] \geq 1 - 1/n$.

**Proof.**

Use union bound:

$$\mathbb{P} [Y_i \geq k$ for some $i] \leq \sum_i \mathbb{P} [Y_i \geq k] \leq 1/n$$

$\square$
Birthday Paradox

Lemma
\[ \Pr[\text{first } m \text{ balls fall in distinct bins}] \leq e^{-m(m-1)/(2n)}. \]

Proof.

- Let \( A_i \) be event that the \( i \)-th ball lands in a bin not containing any of the first \( i - 1 \) balls.
- \[ \Pr[\cap_{1 \leq i \leq m} A_i] = \Pr[A_1] \Pr[A_2|A_1] \ldots \Pr[A_m|\cap_{1 \leq i \leq m-1} A_i] \]
- \[ \Pr[A_i|\cap_{1 \leq i \leq i-1} A_i] = 1 - (i - 1)/n \]
- Putting it together and using \( \sum_{1 \leq i \leq a} i = (a + 1)a/2 \):

\[
\Pr[\cap_{1 \leq i \leq m} A_i] = \prod_{1 \leq i \leq m} \left(1 - \frac{i - 1}{n}\right) \leq e^{-m(m-1)/(2n)}
\]

With \( n = 365 \) and \( m = 29 \), probability \(< e^{-1} \). Tighter analysis is possible.
Coupon Collecting (1/2)

Let $Z_i$ be the throw in which exactly $i$ bins become non-empty. Let $X_i = Z_{i+1} - Z_i$. Note that $Z_n = \sum_{0 \leq i \leq n-1} X_i$

**Lemma**

$\mathbb{E} [Z_n] = nH_n$ where $H_n = 1 + 1/2 + \ldots + 1/n = \ln n + \Theta(n)$.

**Proof.**

- $X_i$ has a geometric distribution:
  
  $$\mathbb{P} [X_i = j] = p_i (1 - p_i)^{j-1}$$

where $p_i = 1 - i/n$.

- $\mathbb{E} [X_i] = 1/p_i$.

- $\mathbb{E} [Z_n] = \sum_{0 \leq i \leq n-1} \mathbb{E} [X_i] = n/n + n/(n - 1) + \ldots + n/1$
Coupon Collecting (2/2)

Let $Z_i$ be the throw in which exactly $i$ bins become non-empty. Let $X_i = Z_{i+1} - Z_i$. Note that $Z_n = \sum_{0 \leq i \leq n-1} X_i$

Lemma
$\mathbb{V}[Z_n] = n^2(\pi^2/6 + o(1)) - nH_n$.

Proof.

- $X_i$ has a geometric distribution: $\mathbb{V}[X_i] = (1 - p_i)/p_i^2$.
- The $X_i$ are independent: $\mathbb{V}[Z_n] = \sum_{0 \leq i \leq n-1} \mathbb{V}[X_i]$
- Therefore, $\mathbb{V}[Z_n]$ equals

$$
\sum_{0 \leq i \leq n-1} \frac{1 - p_i}{p_i^2} = \sum_{0 \leq i \leq n-1} \frac{ni}{(n - i)^2} = n^2 \sum_{1 \leq i \leq n} \frac{1}{i^2} - nH_n
$$

- Appeal to the fact $\lim_{n \to \infty} \left(\sum_{1 \leq i \leq n} i^{-2}\right) = \pi^2/6$. 

\[\square\]