Urn Models

Armaun Sanayei

Stanford University
Department of Computer Science
Department of Mathematics

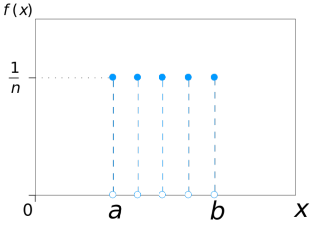
Urn Model

- System of one or more urns containing objects of various types.
  - The objects are called balls.
  - The types are represented by colors or numbers on the balls .
- System evolves in time, subject to rules of throwing balls into the urns or drawing balls under predesignated replacement schemes.
- Drawing a ball is assumed to be completely random.
- Urn Models have many applications.

Realization of Distributions via Urns

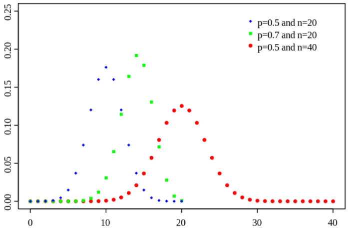
- **Uniform**
- Bernoulli
- **Binomial**
- Negative Binomial
- **Geometric**
- Hyper-geometric
- Negative Hyper-geometric

Discrete Uniform Distribution



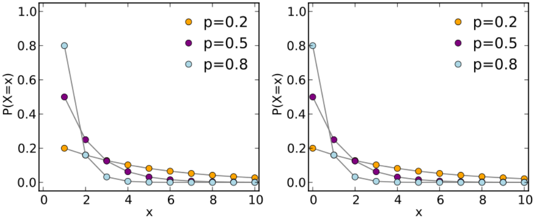
- $P(U = a_k) = \frac{1}{n}$ for k = 1,2, ... n
- Have an urn with n balls labeled with the numbers $a_1, a_2, ..., a_k$. One draw mimics the discrete uniform distribution.

Binomial Distribution



- Given $X \sim$ Binomial(n,p) $\Rightarrow P(X = k) = \binom{n}{k} p^k q^{n-k}$ for k = 1,2, ... n
- Have an urn containing w white balls and b blue balls. Balls are sampled from the urn at random and with replacement. The experiment is carried out n times. The probability p is defined as $\frac{w}{w+b}$. X is the number of white balls chosen.

Geometric



- Given $X \sim \text{Geo(p)} \Rightarrow P(U = a_k) = (1-p)^{k-1}p$ for all positive k
- Have an urn containing w white balls and b blue balls. Balls are sampled from the urn at random and with replacement until a white ball is taken out. The individual drawing has probability $p = \frac{w}{w+b}$. X is the number of draws until the first white ball appears.

Urn Model

Finite-Time Urn Models

- Exact Probabilities
- Combinatorial
- Analytic Solutions

Polya Urn Models

Infinite-Time Urn Models

- Limiting behavior
- Convergence
- Symmetries
- Stochastic Path

Tools for Finite-Time Urn Models

- Recursion
- Conditional Expectation
- Law of Total Probability

Note

These methods often yield formulas that allow us to explicitly understand the system at each discrete time step since the evolution is often deterministic.

Ballot Problems

Problem Statement

Suppose two candidates A and B are running against each other in an election, and A wins (or is tied) by receiving $m \geq n$ votes, where n is the number of votes received by B.

Classical Question: *What is the probability that A stays ahead of B throughout the voting?*

Key Observation - Using Recursion

$$P_{m,n} = \frac{m}{m+n} P_{m-1,n} + \frac{n}{m+n} P_{m,n-1}$$

Solution

We use induction on this recursive relationship to prove:

$$P_{m,n} = \frac{m-n}{m+n}$$

Polya Urn Models

What makes a Polya urn model different from a regular urn model?

Replacement scheme that causes urn to evolve randomly in discrete time steps

Infinite-Time Urn Model

- Limiting behavior
- Convergence
- Symmetries
- Stochastic Path

Example

Start with two balls in an urn. One ball is white the other is black. When you draw out one ball you insert two balls of the same color as the one you drew out.

Tools for Infinite-Time Urn Models

- Ex-changeability $\Rightarrow$ De Finetti's Theorem
- Markov Chains
- Martingale
- Simulations

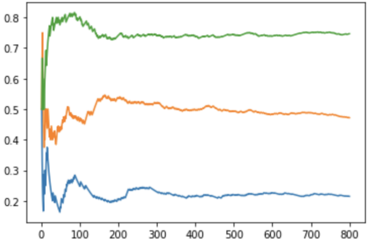
Note

Using these tools often results in understanding of limiting behavior, convergence, and distribution of converged states.

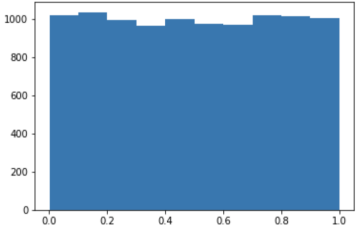
Simulations

Problem

Start with two balls in an urn. One ball is white; the other is black. When you draw out one ball you insert two balls of the same color as the one you drew out.



(a) Convergence of three experiments with same initial conditions



(b) Limit points of 10000 trials

Figure: Simulation Results

Tools for Approaching Infinite-time Urn Models

Ex-changeability

A random sequence $\{E_n\}_{n\geq 1}$ with $E_n = 0$ or $E_n = 1$ is called exchangeable if for any finite n and any permutation $\sigma \in S_n$

$$(E_1, \ldots, E_n) \stackrel{d}{=} \left(E_{\sigma(1)}, \ldots, E_{\sigma(n)}\right)$$

Ex-changeability gives us access to **De Finetti's Theorem**

Martingale

A discrete-time random process $X_{n_{n\geq 1}}$ is called a martingale if, for all n,

$$E[X_{n+1}|X_1, ..., X_n] = X_n.$$

Markov Chains

Discussed Later

The Ehrenfest Urn

Problem Statement

Suppose that we have two urns, labeled 0 an 1, that contain a total of m balls. The state of the system at time n is the number of balls in urn 1, which we will denote by X_n.

Approach

- Determine whether the urn is Markov, Ex-changeable, or a Martingale.
- Determine transition probability based on last time step (given that it is Markov)
- Use the matrix to understand dynamic behavior

Application

The Ehrenfest urn, like many urn models, has many applications. One of the most well-know application for this urn is for the exchange of gas molecules between two containers.

Markov-Chain

A discrete-time random process $X_{n n\geq 1}$ taking values in $\mathbf{Z}$ is called a Markov Chain if for all n and $a \in \mathbf{Z}$,

$$P(X_{n+1} = a | X_1, ..., X_n) = P(X_{n+1} = a | X_n).$$

This relationship is often represented in a transition matrix where each column and row represents a state in the state space.

$$P = \begin{bmatrix} P_{1,1} & P_{1,2} & \dots & P_{1,j} & \dots & P_{1,S} \\ P_{2,1} & P_{2,2} & \dots & P_{2,j} & \dots & P_{2,S} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ P_{i,1} & P_{i,2} & \dots & P_{i,j} & \dots & P_{i,S} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ P_{S,1} & P_{S,2} & \dots & P_{S,j} & \dots & P_{S,S} \end{bmatrix}$$

Ehrenfest Urn Solution

Key Observation

$$P_{i,j} = \begin{cases} \frac{i}{M} & , \quad j = i - 1 \\ \frac{M-i}{M} & , \quad j = i + 1 \\ 0 & , \quad \text{otherwise} \end{cases}$$

This translates to a Markov chain transition matrix where the i,j entry is the $P_{i,j}$ value from the above transition probability rule.