A Stronger Concentration

Theorem 1. *Let* $r = m/n$. *Then for any* $\lambda > 0$,

$$\Pr(|Z - \mu| \geq \lambda) \leq 2e^{-\frac{\lambda^2(n-1/2)}{n^2-\mu^2}}.$$

Proof:

We bound $|Z_t - Z_{t-1}|$ more carefully.

Define $z(Y,t)$ as the expectation of Z given that Y bins are empty at time t.

$z(Y,t) = E[Z|Y$ bins are empty at time $t]$

$= Y(1-1/n)^{m-t}$.

Let Y_t denote the number of empty bins at time t.

$Z_{t-1} = z(Y_{t-1}, t-1) = Y_{t-1}(1-1/n)^{m-t+1}$.

At time t there are two possibilities:

1. The tth ball goes into a currently non-empty bin.

Then $Y_t = Y_{t-1}$, and

$Z_t = z(Y_t, t) = z(Y_{t-1}, t) = Y_{t-1}(1 - 1/n)^{m-t}$.

2. The t ball goes into a currently empty bin.

Then $Y_t = Y_{t-1} - 1$ and

$Z_t = z(Y_t, t) = z(Y_{t-1} - 1, t) = (Y_{t-1} - 1)(1 - 1/n)^{m-t}$.

We bound $\Delta_t = Z_t - Z_{t-1}$.

1. With probability $1 - Y_{t-1}/n$, Δ_t is

$Y_{t-1}(1 - 1/n)^{m-t} - Y_{t-1}(1 - 1/n)^{m-t+1}$

$= \frac{Y_{t-1}}{n}(1 - 1/n)^{m-t}$

2. With probability Y_{t-1}/n, Δ_t is

$(Y_{t-1} - 1)(1 - 1/n)^{m-t} - Y_{t-1}(1 - 1/n)^{m-t+1}$

$= -(1 - \frac{Y_{t-1}}{n})(1 - 1/n)^{m-t}$

Thus,

$-(1 - 1/n)^{m-t} \leq \Delta_t \leq (1 - 1/n)^{m-t}$.

Thus $|Z_t - Z_{t-1}| \leq c_t$ where $c_t = (1 - 1/n)^{m-t}$.

$\sum_{t=1}^{m} c_t^2 = \frac{n^2 - \mu^2}{2n-1}$

Azuma gives the result.

Markov Chains

Let $\mathcal{S}$ be a set of states.

A stochastic process $\mathbf{X} = \{X_t : t \in T\}$ is a collection of random variables.

If for all t, X_t assumes values from a countably infinite set, then $\mathbf{X}$ is a *discrete space* process.

If X_t assumes values from a finite set, then the process is **finite**.

If T is a countably infinite set we say that $\mathbf{X}$ is a discrete time process.

Definition 1. *A discrete time stochastic process* $X_0, X_1, X_2, ...$ *is a Markov chain if*

$$Pr(X_i \mid X_0, X_1, X_2, ...) = Pr(X_i \mid X_{i-1}).$$

Transition Matrix

A Markov chain is fully defined by a **transition matrix** $\mathcal{P}$ where for all $i, j \in \mathcal{S}$, $P_{i,j}$ gives the probability of moving from state i to state j (in one step).

$$P_{i,j} = \Pr(X_t = j | X_{t-1} = i).$$

Graph Representation:

A Markov chain can be represented by a directed, weighted graph.

The vertices represent the states and the edges represent the transitions between states.

Each outgoing edge has a probability (sum over outgoing edges for each vertex is 1).

The process chooses the next transition according to the probabilities of the outgoing edges.

Example

Consider a system with a total of m balls in two containers.

We start with all balls in the first container.

At each step we choose a ball uniformly at random from all the balls and with probability 1/2 move it to the other container.

Let X_i denote the number of balls in the first container at time i.

$X_0, X_1, X_2, ...$ defines a Markov chain with the following transition matrix:

$$p_{i,j} = \begin{cases} \frac{m-i}{2m} & j = i+1 \\ \frac{i}{2m} & j = i-1 \\ \frac{1}{2} & j = i \\ 0 & |i-j| > 1 \end{cases}$$

m-step transitions

For $m \geq 0$, we define the m-step transition probability

$$P^m_{i,j} = \Pr(X_{t+m} = j | X_t = i)$$

as the probability that the chain moves from state i to state j in exactly m steps.

$P^m_{i,j} = \sum_{k \geq 0} P_{i,k} P^{m-1}_{k,j}$.

Let $\mathcal{P}^{(m)}$ be the matrix whose entries are the m-step transition probabilities.

Then, $\mathcal{P}^{(m)} = \mathcal{P}.\mathcal{P}^{(m-1)} = \mathcal{P}^m$.

Let $p_i(t)$ denote the probability that the Markov chain is at state i at time t.

Let $\hat{p}(t) = (p_0(t), p_1(t), p_2(t), \ldots)$ be the vector giving the state distribution of the chain at time t.

Summing over all possible states at time $t-1$, we have

$$p_i(t) = \sum_{j \geq 0} p_j(t-1) P_{j,i}$$

Or, $\hat{p}(t) = \hat{p}(t-1)\mathcal{P}$.