Header styling inspired by CS 70: https://www.eecs70.org/

1 From Last Time

- The reason we need to divide by 400 is that CLT states that the sample average $\frac{S_n}{n}$ looks like a normal distribution with mean μ and standard deviation $\frac{\sigma^2}{\sqrt{n}}$
- Then, if we get rid of the mean by subtracting by μ and dividing by $\frac{\sigma^2}{\sqrt{n}}$, we a standard normal distribution.
- So, the equation looks like:
$$\frac{\frac{S_n}{n} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$
Then, the way that we say that it looks like the standard normal is to say that the PDF for the distribution of $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ looks like the PDF for a standard normal distribution.

2 Markov Chains

- Mathematically, we deal with a random variable X_n, which describes the state of the markov chain at time n.
- At initial time 0, we are given an initial distribution vector π_0 such that
$$P(X_0 = 0) = \pi_0(0) \quad P(X_0 = 1) = \pi_0(1)$$
Then, the transitions are governed by the probabilities:
$$P(X_{n+1} = 1 \mid X_n = 0) = 1 - a$$
- Amnesic property: the probabilities don't change based on the path taken to get to the stage. This is basically the analogue of independence.
- You can also represent the markov chain in linear algebra language:
$$P(i, j) = P(X_{n+1} = j \mid X_n = i)$$
P is then called the *transition matrix* for the markov chain.
- Given a initial distribution π_0, then we can write $\pi_n = \pi_0 P^n$, since each time step is given by multiplication on the left by P, so after n time steps we just scale that by n. A lot of the time, π_0 is concentrated at a single state.

2.1 Invariant Distributions

- A distribution is invaraint if $\pi_{n+1} = \pi_n P$. Then, we have $\pi_n = \pi_0$ if and only if π_0 is invariant.
- A markov chain is irreducible if it can go from every state i to every other state j in a finite number of steps.
- A Markov chain is aperiodic if the greatest common divisor of the set of path lenghts from any two nodes i and j is 1.
- Fundamental theorem of Markov Chains: For any finite, irreducible Markov Chain, the probability distribution at time n for any initial state X_0 converges to π, where π is the unique invariant distribution for the markov chain. So, for any X_0 and any state i, we have $P(X_n = i) = \pi(i)$ as $n \to \infty$.