The specific examples I’m discussing are not so important; the point of these first lectures is to illustrate a few of the 100 ideas from STAT134.

**Ideas used in Lecture 4.**

- Conditional expectation as a random variable.
- Uses of $\mathbb{E}[\mathbb{E}(X|Y)] = \mathbb{E}X$.
- Uniform random point in a region.
- If $X$ has continuous distribution function $F$ then $F(X)$ has uniform distribution on $(0, 1)$.
- Conditioning on first step, for simple symmetric random walk.
A Markov chain \((X_0, X_1, X_2, \ldots) = (X_t, t \geq 0)\) is a process such that

(i) each \(X_t\) takes values in the same state space \textbf{States}

(ii) There are numbers \((p_{ij}, i, j \in \textbf{States})\) such that

\[
P(X_{t+1} = j | X_t = i, X_{t-1} = i_{t-1}, \ldots X_0 = i_0) = p_{ij}
\]

for all \(t, i, j\) and all \((i_0, \ldots, i_{t-1})\).

In words, (ii) says that at each time \(t\), probabilities for the future depend on the current state \(X_t\) but not on past states.

We can consider the matrix \(P\) with entries \((p_{ij})\). For the definition to make sense, \(P\) must have the properties

(iii) \(p_{ij} \geq 0\), for all \(i, j\).

(iv) \(\sum_j p_{ij} = 1\), for all \(i\).

A matrix with these properties is called a \textbf{stochastic matrix}. It is intuitively clear that, given any stochastic matrix \(P\) indexed by \textbf{States}, there exists the Markov chain specified by (i,ii).
So for a Markov chain \((X_0, X_1, X_2, \ldots)\)

(ii) There are numbers \((p_{ij}, i, j \in \text{States})\) such that

\[
P(X_{t+1} = j | X_t = i, X_{t-1} = i_{t-1}, \ldots X_0 = i_0) = p_{ij}
\]

for all \(t, i, j\) and all \((i_0, \ldots, i_{t-1})\). In this context we call \(P = (p_{ij})\) the \textbf{transition matrix} and call the \(p_{ij}\) the \textbf{transition probabilities} for the chain.

If we want to calculate a probability or expectation for a Markov chain, the answer will depend not only on \(P\) but also on the “initial distribution” of \(X_0\). Often we think of the initial state as non-random: \(X_0 = i_0\).
We can visualize $\mathbf{P}$ as a weighted directed graph; draw edge $i \to j$ if $p_{ij} > 0$ and assign “weight” $p_{ij}$ to that edge. Then visualize the chain as a jumping particle; from present state $i$ the particle will, at the next step, jump to state $j$ with probability $p_{ij}$.

Textbook [PK] sections 3.1-3.2 gives numerical examples of matrices with 3 or 4 states. You should read this. I do not emphasize numerics, but will do one example on the board.

I will give 5 examples – meaning an explicit set of States and an explicit transition matrix $\mathbf{P}$. Some of these are “toy models”, meaning we are imagining some real-world process but making a hugely over-simplified and unrealistic model. Most of the examples are in [PK] section 3.3.
Example. Recall simple symmetric random walk

\[ X_t = \sum_{i=1}^{t} \xi_i \]

where \((\xi_i)\) are i.i.d. with \(\mathbb{P}(\xi_i = 1) = \mathbb{P}(\xi_i = -1) = \frac{1}{2}\).

Here \((X_t)\) is the Markov chain with States = \(\mathbb{Z}\) and

\[ (*) \quad p_{i,i-1} = \frac{1}{2}, \quad p_{i,i+1} = \frac{1}{2}. \]

In the “gambler’s ruin” variant, where you stop on reaching \(K\) or 0, we take the states as \(\{0, 1, 2, \ldots, K\}\) and modify (*) by setting

\[ p_{00} = 1, \quad p_{KK} = 1. \]

Note the implicit convention: if \(p_{ij}\) is not specified then \(p_{ij} = 0\).
Example: Ehrenfest urn model.  
2 boxes, $2a$ balls, each ball in one of the boxes. Each step, pick uniform random ball and move to other box.

Consider $Y_t =$ number of balls in left box after $t$ steps,  
**States** = \{0, 1, 2, \ldots, 2a\}.

\[
p_{i,i-1} = \frac{i}{2a}, \quad p_{i,i+1} = \frac{2a - i}{2a}.
\]
Example: Fisher-Wright genetic model. (2-type, no mutation or selection).

- $2N$ genes in each generation, of types $a$ or $A$.
- “children choose parents”: each gene is a copy (same type) of a uniform random gene from previous generation.

Then

$$X_t = \text{number of type-}a \text{ in generation } t$$

is a Markov chain, with states \{0, 1, 2, \ldots, 2N\} and transition probabilities

$$p_{ij} = \mathbb{P}(\text{Bin}(2N, \frac{i}{2N}) = j) = \binom{2N}{j} \left(\frac{i}{2N}\right)^j \left(\frac{2N-i}{2N}\right)^{2N-j}.$$
Queue models are more naturally set up in continuous time, but here is a **Discrete time queue model**.

- Service takes unit time for each customer.
- If no customer, server takes a break for unit time.
- \( \xi_t \) new customers arrive during time \([t - 1, t]\).
- Model \((\xi_1, \xi_2, \ldots)\) as i.i.d.

Consider

\[
X_t = \text{number of customers at time } t.
\]

Clearly

\[
X_t = (X_{t-1} - 1)^+ + \xi_t.
\]

Here \((X_t)\) is a Markov chain on states \(\{0, 1, 2, \ldots\} = \mathbb{Z}^+\) with transition probabilities

\[
p_{0j} = \mathbb{P}(\xi = j), j \geq 0
\]

\[
p_{ij} = \mathbb{P}(\xi = j - i + 1), i \geq 1, j \geq i - 1.
\]
Example: Umbrellas.

- A man owns $K$ umbrellas, which are either at home or at work.
- He goes to work each morning, and goes home each evening.
- If raining, he takes an umbrella, if one is available. If not raining he does not take an umbrella.
- Model (unrealistic) that $\mathbb{P}(\text{rain}) = p$, independently, each morning and evening.

To set up as a Markov chain, consider

$$X_t = \text{number of umbrellas at home, end of day } t.$$ 

States $\{0, 1, \ldots, K\}$.

What are the transition rates?
Example: Umbrellas.

- A man owns $K$ umbrellas, which are either at home or at work.
- He goes to work each morning, and goes home each evening.
- If raining, he takes an umbrella, if one is available. If not raining he does not take an umbrella.
- Model (unrealistic) that $\mathbb{P}(\text{ rain }) = p$, independently, each morning and evening.

To set up as a Markov chain, consider

$$X_t = \text{ number of umbrellas at home, end of day } t.$$ 

States $\{0, 1, \ldots, K\}$.

$p_{01} = p, \quad p_{00} = 1 - p$

$p_{K,K-1} = p(1 - p), \quad p_{KK} = 1 - p(1 - p)$

$p_{i,i+1} = p_{i,i-1} = p(1 - p), \quad p_{ii} = 1 - 2p(1 - p), \quad 1 \leq i \leq K - 1.$
Conceptual point. The notion of independence is used in two conceptually different ways.

- We often use independence as an assumption in a model – throwing dice, for instance.
- Given a well-defined math model, events or random variables $X, Y$ either are independent, or are not independent, as a mathematical conclusion.

We see the same point in these examples of Markov chains. For “simple symmetric random walk” and “discrete time queue model” we started with a model defined using i.i.d. random variables, then defined $X_t$ in terms of that model. In the other examples we started with a story in words, and then built a math model which assumed the Markov property.