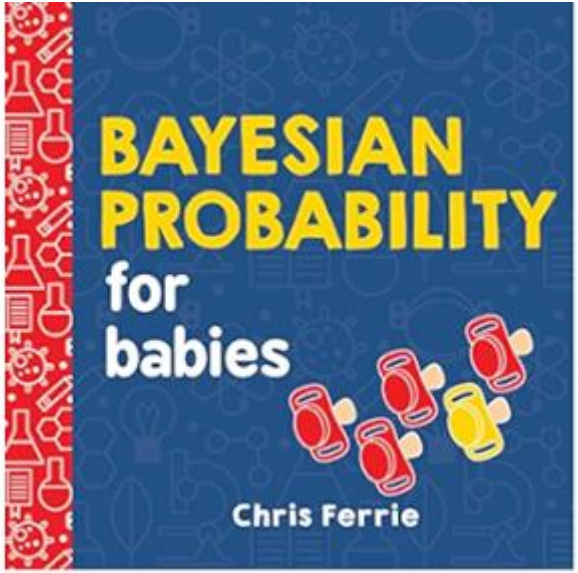
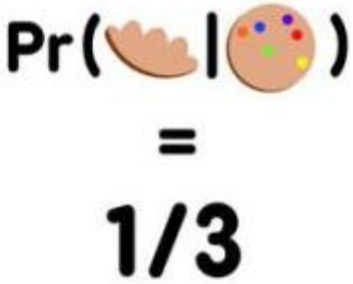
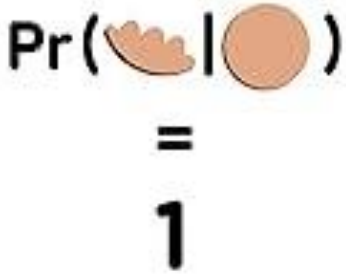


If the cookie had candy, then very few bites would have no candy.



The probability of a no-candy bite, given a candy cookie, is 1/3.

If the cookie had no candy, then every bite would have no candy.



The probability of a no-candy bite, given a no-candy cookie, is 1.

CS 1671/2071
Human Language Technologies

Session 3: Linear algebra, probability review

Michael Miller Yoder

January 15, 2025

University of Pittsburgh | School of Computing and Information

Overview: Linear algebra and probability review

1. Course logistics
2. JupyterHub setup and preprocessing activity
3. Probability review
4. Linear algebra review

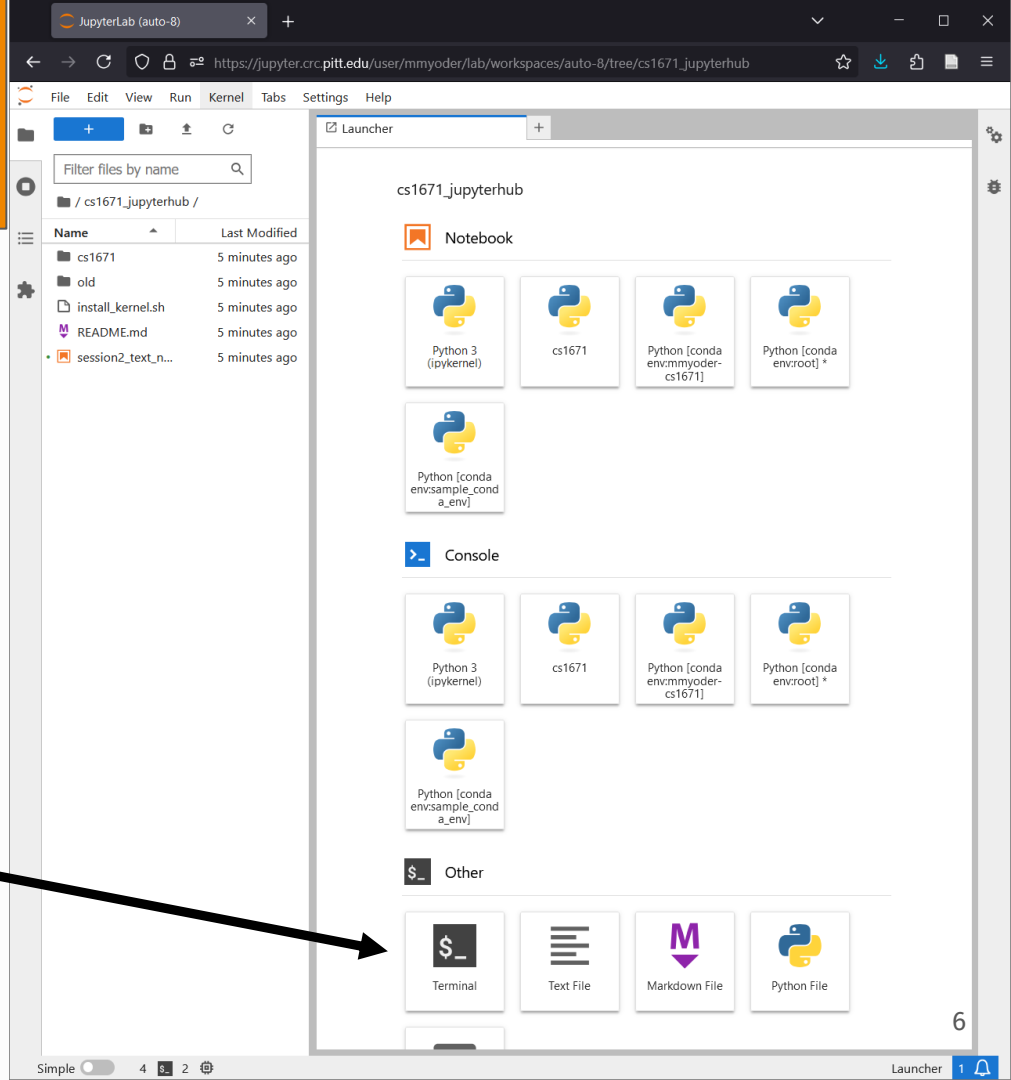
Course logistics

- No class next Mon for MLK Day
- Next class is next Wed Jan 22
- Homework 1 is due next Thu Jan 23

JupyterHub setup and activity

Set up Python virtual environment

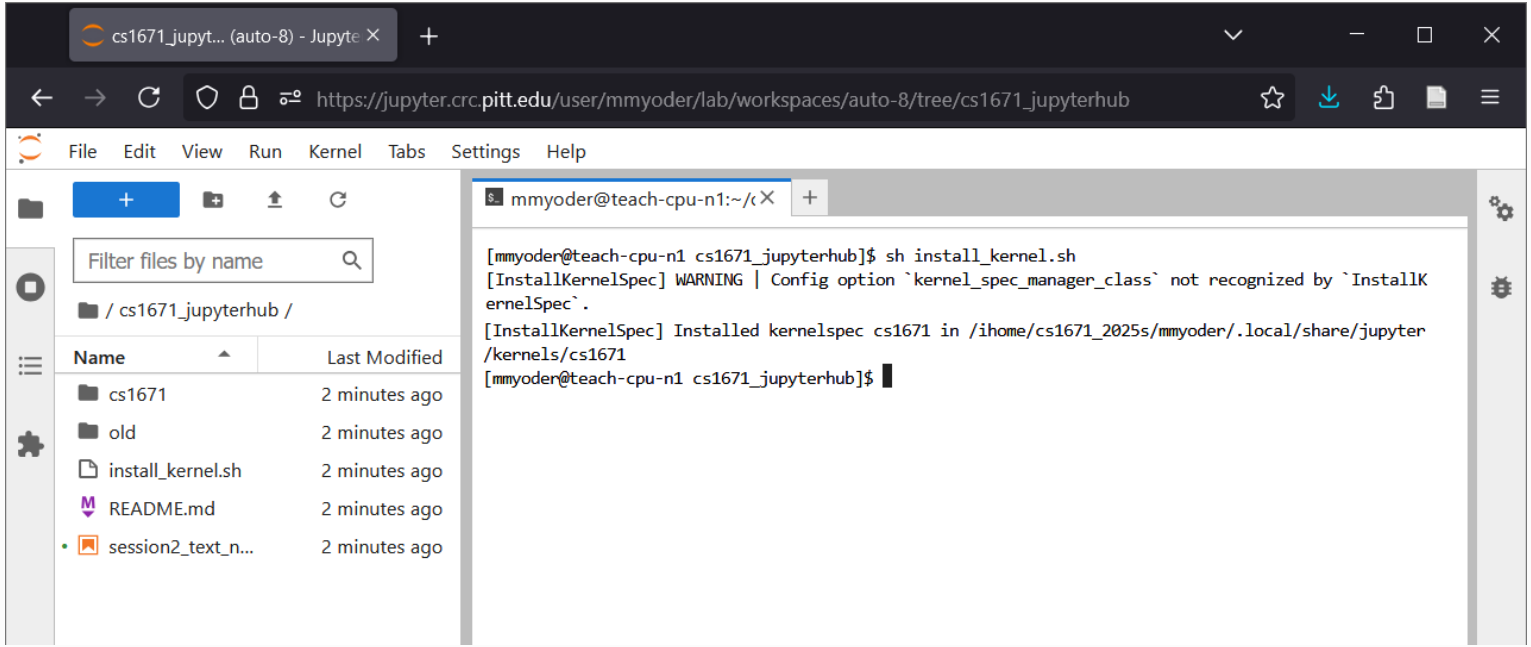
1. Go to this nbgitpuller link
   - Log in with your Pitt username
   - Start a server with **Teach – 6 cores, 3 hours**
   - This should pull a folder (cs1671_jupyterhub) into your JupyterLab
2. Open a terminal



Set up Python virtual environment

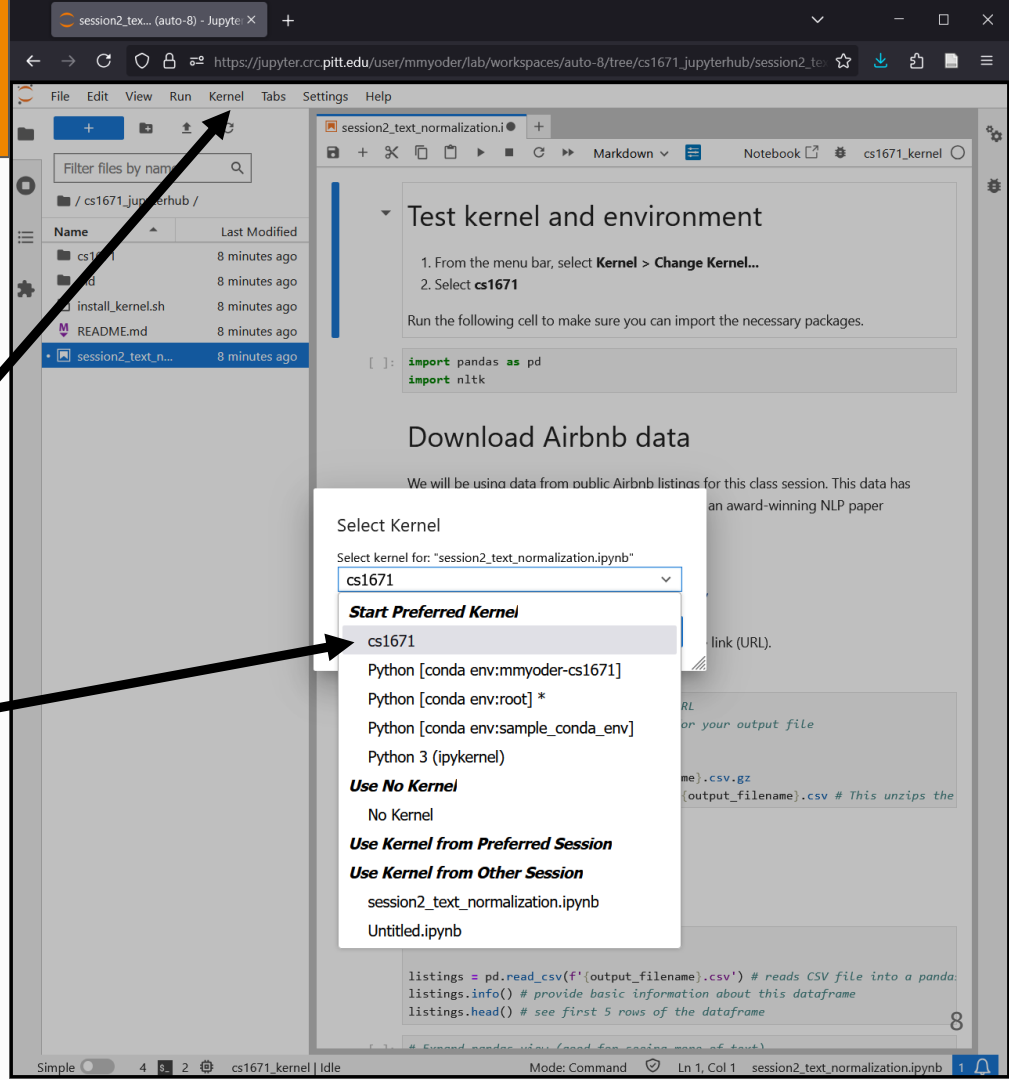
In a terminal, run

```
sh install_kernel.sh
```



Open Jupyter Notebook

1. Double-click **session2_text_normalization .ipynb** on the left panel to open the notebook
2. From the top menu, click **Kernel > Change Kernel...**
3. Select **cs1671** as your kernel
4. Run the first code cell under **Test kernel and environment** that imports pandas and nltk

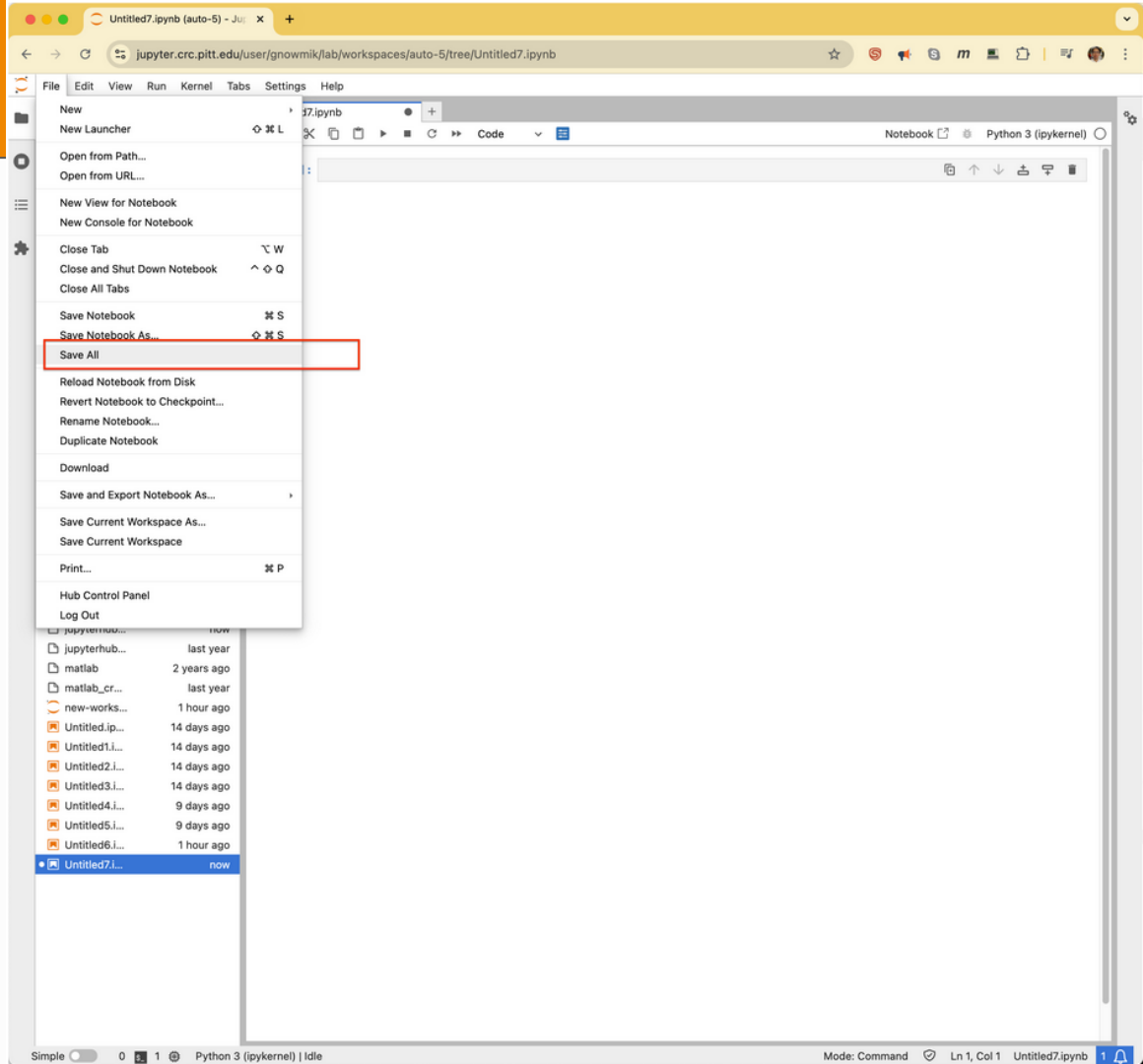


Preprocessing Airbnb listings

Implementation

- Remove undesired text with regular expressions
- Lowercase
- Remove stopwords
- Tokenize with the NLTK package
- Stem the tokens with NLTK

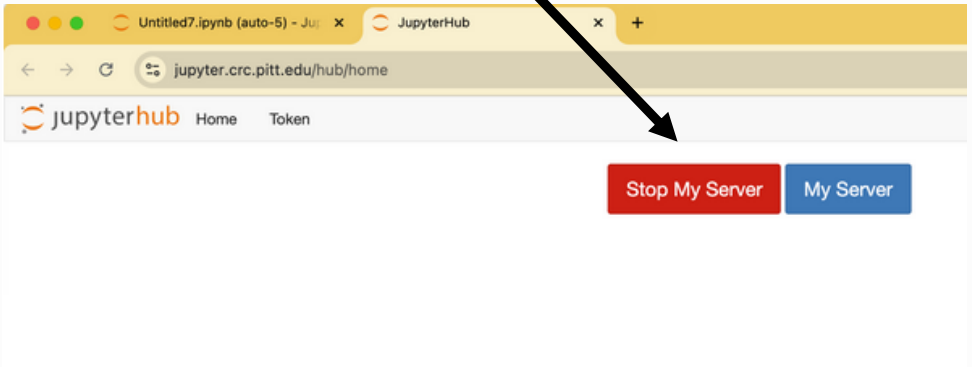
Saving your work

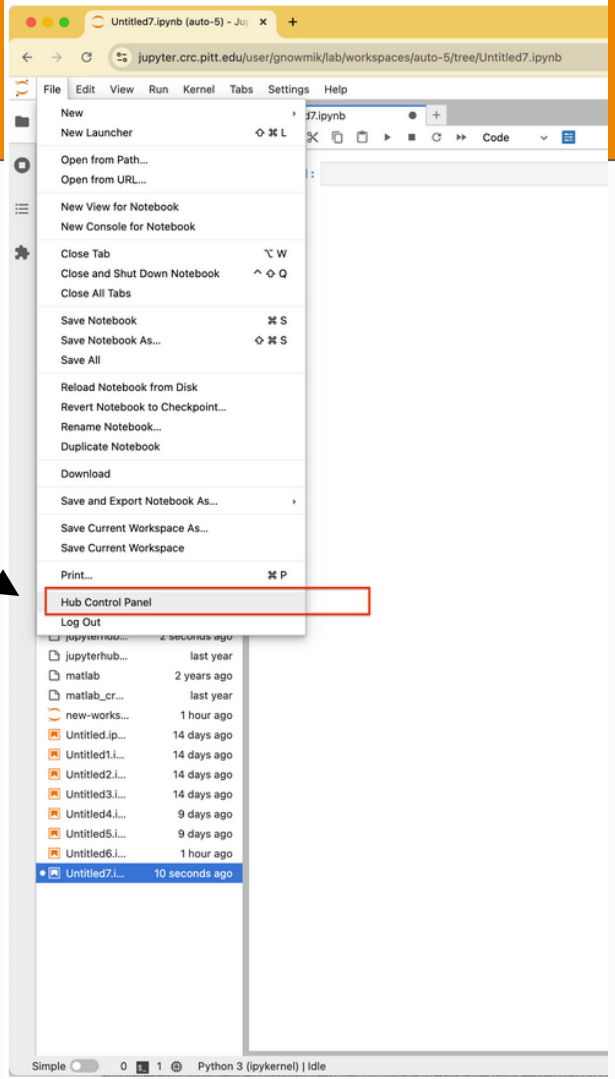


Ending your session

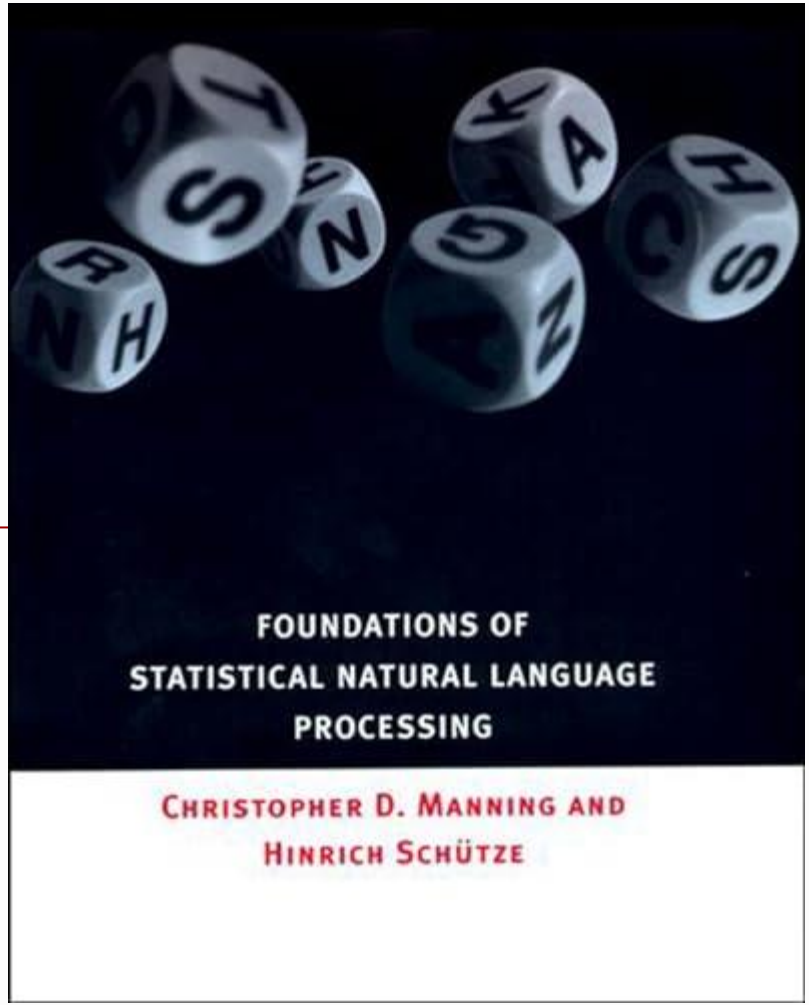
Be sure to save your work before ending the session

1. Select **File** > **Hub Control Panel**
2. Click **Stop My Server**





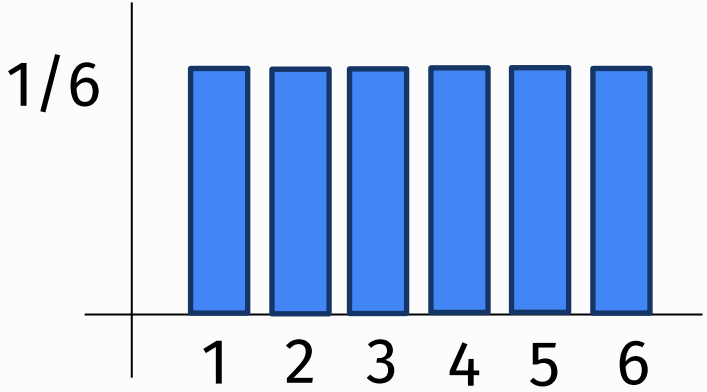
Probability review

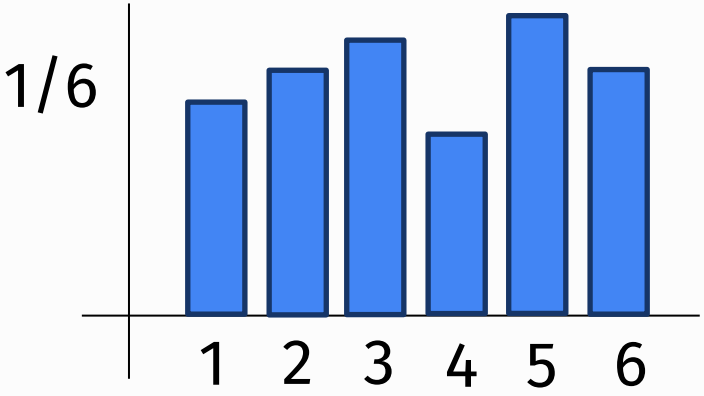


Probability

- Probability of an event a occurring
- $P(a)$
  - For example, a could be a die showing a 2 out of $\{1, 2, 3, 4, 5, 6\}$
- Estimate $P(a)$ as $\dfrac{\text{count}(a)}{\text{count}(\textit{all events})}$
  - Relative frequency or maximum likelihood estimate (MLE)

Probability distributions





Random variables

- **Random variable:** a mapping from a domain of possible outcomes in a sample space to a range of measurable space, such as counts
  - Typically the “result of an experiment”
  - For example, flipping a coin multiple times (possible outcomes {H, T}) and recording the result as 0 for tails and 1 for heads
- Distribution of a random variable *X*
  - *P(X)* is a probability distribution over all possible values in the sample space. Probability mass function
  - *P(X = x)* is the probability that the random variable *X* has the value *x*
  - *P(X = heads)*, where *X* is the random variable of a coin flip

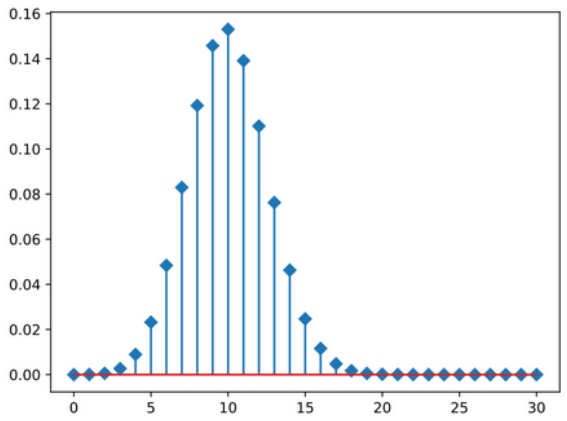


Figure 7.1: $\mathbf{P}(k \text{ heads})$ in 30 tosses, success prob 1/3.

Figure from Baxter 2023

Joint probability

- Probability of 2 events both occurring

$$P(A \cap B)$$

$$P(A, B)$$

- When rolling 2 dice, what's the probability of getting two 5s?

Let D_1 be dice 1, D_2 be dice 2. These events are independent, so:

$$P(D_1 = 5, D_2 = 5) = P(D_1 = 5) \cdot P(D_2 = 5)$$

$\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$ since there are 36 different possible combinations

Conditional probability

- Probability distributions sometimes change if you know another event has occurred or not occurred
- **Conditional** probability of an event a occurring **given that another event, b, has already occurred**
  - $P(a|b)$
- Assume
  - X is the outcome of rolling a die once
  - F is the event $X = 6$
  - E is the event $X > 4$
- Die is rolled and we are told that E has occurred
- What is $P(F|E)$?

Adapted from Roussas 2007

Conditional probability

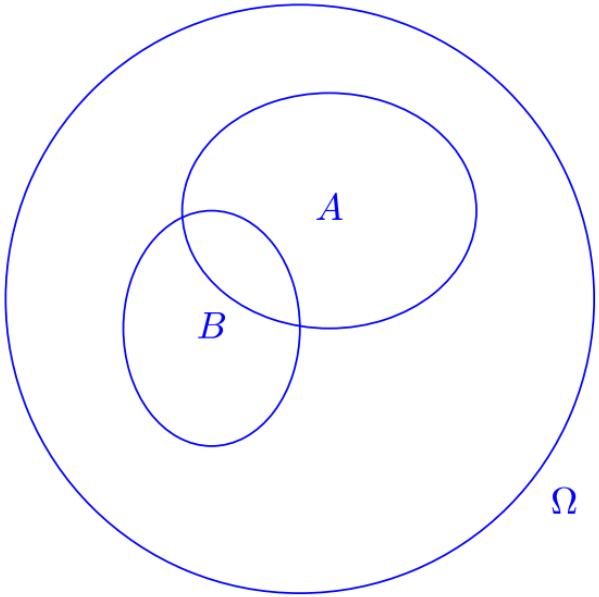


Figure 4.1: Events on the dart board

- Assume a very bad dart thrower (maybe Michael)

$$\mathbf{P}(A) = \frac{\mathbf{area}(A)}{\mathbf{area}(\Omega)}$$

Adapted from Baxter 2023

Conditional probability

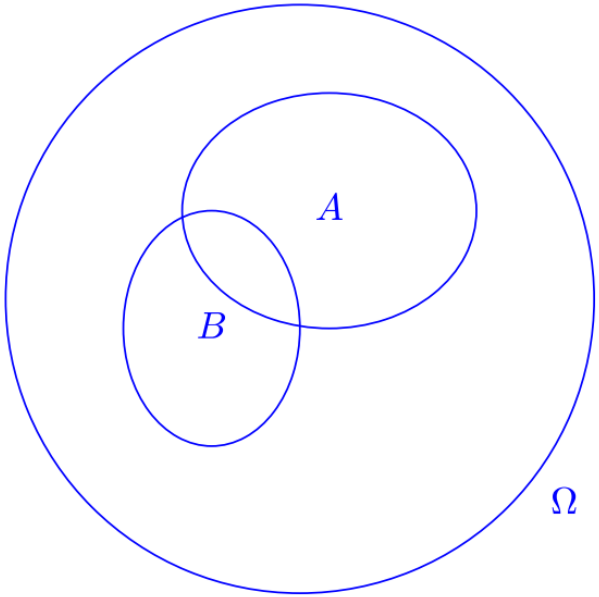


Figure 4.1: Events on the dart board

- You don't see the throw, but somebody tells you that the dart landed in *B* (so *B* occurred)
- What is the formula for P(*A*|B)?

Adapted from Baxter 2023

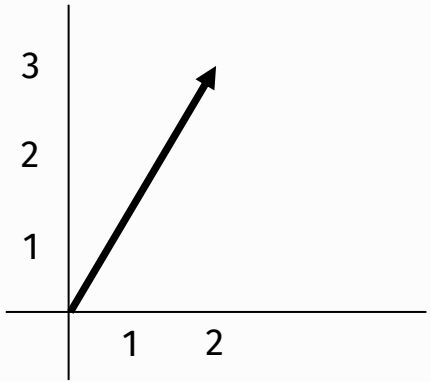
Linear algebra review

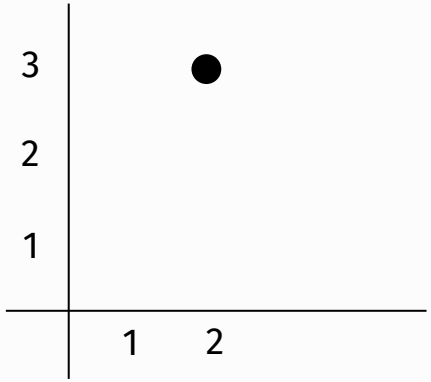
Vectors

An array of numbers with D dimensions

$$[\ 2\ 3\ ]$$

Can be represented as a point in D-dimensional space





Dot product: vector · vector

Sum of the products of each vector dimension

$$\boldsymbol{v} \begin{bmatrix} v_1 & v_2 & \cdots & v_N \end{bmatrix} \cdot \boldsymbol{w} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix}$$

$$\boldsymbol{v} \cdot \boldsymbol{w} = \sum_{i=1}^{N} v_i w_i = v_1 w_1 + v_2 w_2 + \cdots + v_N w_N$$

Matrices

A matrix is an array of numbers

$$\begin{bmatrix} 6 & 4 & 24 \\ 1 & -9 & 8 \end{bmatrix}$$

Two rows, three columns.

Slide adapted from David Mortensen

It's Easy to Multiple a Matrix by a Scalar

$$2 \cdot \begin{bmatrix} 5 & 2 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 2 \cdot 5 & 2 \cdot 2 \\ 2 \cdot 3 & 2 \cdot 1 \end{bmatrix} = \begin{bmatrix} 10 & 4 \\ 6 & 2 \end{bmatrix}$$

Slide adapted from David Mortensen

Dot product: vector · matrix

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \quad \\ \quad \\ \quad \end{bmatrix}$$

Dot product: matrix · matrix

Let a_1 and a_2 be the row vectors of matrix A and b_1 and b_2 be the column vectors of a matrix B. Find $C = AB$

$$\begin{bmatrix} 1 & 7 \\ 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} 3 & 3 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} a_1 \cdot b_1 & a_1 \cdot b_2 \\ a_2 \cdot b_1 & a_2 \cdot b_2 \end{bmatrix} = \begin{bmatrix} 38 & 17 \\ 26 & 14 \end{bmatrix}$$

A must have the same number of rows as B has columns.

Slide adapted from David Mortensen

Questions?

No class next Mon for MLK Day.
Will see you again on Wed.
Take a look at HW1