Slides with solutions

After class each day we will post copies of the slides with solutions to (most of) the in class problems.
Welcome to 18.05
Introduction to Probability and Statistics
Spring 2016

A weighted random number generator just produced a new batch of numbers.

Let's use them to build narratives!

All sports commentary

http://xkcd.com/904/
Staff

- **David Vogan**
  dav@math.mit.edu, office hours Friday 10–12 in 2-355

- **Sam Hopkins**
  shopkins@mit.edu, office hours *TBA in room TBA*

- **Gus Lonergan**
  gusl@mit.edu, office hours *TBA room TBA*

- **Lucas Mason-Brown**
  lmbrown@mit.edu, office hours *TBA room TBA*
Course materials are at:
http://math.mit.edu/~dav/05.dir/05.html

There is an 18.05 Stellar site, but for now we’re not using it.

Site will have all reading materials and problem sets.

Copies of the slides with solutions to all problems discussed in class will be posted after each class.
Active Learning

Read the ‘general information’ slide on the web site.

Before class

- Reading and reading questions.
- Reading questions count toward grade.
- Lecture will assume you’ve done the reading.

In class:

- Combination of lecture and problem solving
- We won’t assume you’ve completely mastered the reading.
  - We will assume you’ve read the reading.
  - Use the Piazza discussion board—link is on the web site.
  - Bring questions to class.
Class

Class Time
- MW: Lecture/clicker questions/board questions
  Participation on clicker questions counts towards your grade
  No computer use in class on TR.
- F: **R Studio**—bring your laptop

In-class Groups
- Groups of 3.
- You will be able to choose your own group.
- If you need to find a group or your group needs a third person let us know and we’ll help.

**R:** for computation, simulation and visualization
- will teach you everything you need
- no hardcore programming.
Problem Sets

- Usually due on Mondays
- Turn in to the 18.05 slots outside 4-174 by 9:30 AM
- You’ll eventually be able to check your numerical answers to problems on the web site before the due date. (This may not be running for the first problem set.)
- Problem sets will be graded on the logic and explanation of your answer.
Clickers, R, Piazza

Clickers

- Buy a TurningTechnology clicker at MIT coop.
- Follow the instructions for registering it to this class in the 'general information' section of the web site.

R

- Free open source package.
- Very easy to use and install.
- Instructions and a link for this are on the web site.

Piazza

- We will use the Piazza discussion forum.
- Mostly for students to ask questions of each other.
- Sign up by following the link from our MITx site.
Everything we just went over and more is **on the website!**  _rtfw!_
For Next Time

- Familiarize yourself with the web site
- Get and register clicker
- Install R and R Studio

- Sign up for Piazza and join our class. (Link on the web site)
- Read class 1 notes (summary of what we’ll do today)
- Go through the class 2 prep pages and answer the reading questions
Platonic Dice
Probability vs. Statistics

Different subjects: both about random processes

Probability
- Logically self-contained
- A few rules for computing probabilities
- One correct answer

Statistics
- Messier and more of an art
- Get experimental data and try to draw probabilistic conclusions
- No single correct answer
What is the probability of getting exactly 1 heads in 3 tosses of a fair coin?
Poker Hands

Deck of 52 cards

- 13 ranks: 2, 3, ..., 9, 10, J, Q, K, A
- 4 suits: ♥, ♠, ♦, ♣,

Poker hands

- Consists of 5 cards
- A one-pair hand consists of two cards having one rank and the remaining three cards having three other rank
- Example: \{2♥, 2♠, 5♥, 8♣, K♦\}

The probability of a one-pair hand is:
1. less than 5%
2. between 5% and 10%
3. between 10% and 20%
4. between 20% and 40%
5. greater than 40%
Sets in Words

Old New England rule: don’t eat clams (or any shellfish) in months without an ’r’ in their name.

- $S =$ all months
- $L =$ the month has 31 days
- $R =$ the month has an ‘r’ in its name

$S = \{\text{Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec}\}$

$L = \{\text{Jan, Mar, May, Jul, Aug, Oct, Dec}\}$

$R = \{\text{Jan, Feb, Mar, Apr, Sep, Oct, Nov, Dec}\}$

$L \cap R = \{\text{Jan, Mar, Oct, Dec}\} =$ months with 31 days and an ‘r’
Visualize Set Operations with Venn Diagrams

- $S$
- $L$
- $R$
- $L \cup R$
- $L \cap R$
- $L^c$
- $L - R$
Product of Sets

\[ S \times T = \{(s, t)\} \]
Inclusion-Exclusion Principle

$S$

$A \cap B$

$B$

$A$
A band consists of singers and guitar players.

- 7 people sing
- 4 play guitar
- 2 do both

How many people are in the band?
Rule of Product

3 shirts, 4 pants = 12 outfits

(More powerful than it seems.)
Concept Question: DNA

DNA is made of sequences of nucleotides: A, C, G, T.

How many DNA sequences of length 3 are there?
(i) 12  (ii) 24  (iii) 64  (iv) 81

answer: (iii) $4 \times 4 \times 4 = 64$

How many DNA sequences of length 3 are there with no repeats?
(i) 12  (ii) 24  (iii) 64  (iv) 81

answer: (ii) $4 \times 3 \times 2 = 24$
Board Question 1

There are 5 Competitors in 100m final.

How many ways can gold, silver, and bronze be awarded?

answer: $5 \times 4 \times 3$.

There are 5 ways to pick the winner. Once the winner is chosen there are 4 ways to pick second place and then 3 ways to pick third place.
Board Question 2

I won’t wear green and red together; I think black or denim goes with anything; Here is my wardrobe.

Shirts: 3B, 3R, 2G; sweaters 1B, 2R, 1G; pants 2D,2B.

How many different outfits can I wear?
**Solution**

**answer:** Suppose we choose shirts first. Depending on whether we choose red compatible or green compatible shirts there are different numbers of sweaters we can choose next. So we split the problem up before using the rule of product. A multiplication tree is an easy way to present the answer.

![Multiplication Tree Diagram]

Multiplying down the paths of the tree:
Number of outfits = \((3 \times 3 \times 4) + (3 \times 4 \times 4) + (2 \times 2 \times 4) = 100\)
Permutations

Lining things up. How many ways can you do it?

‘abc’ and ‘cab’ are different permutations of \{a, b, c\}
Permutations of $k$ from a set of $n$

Give all permutations of 3 things out of $\{a, b, c, d\}$
Permutations of $k$ from a set of $n$

Give all permutations of 3 things out of \{a, b, c, d\}

\[
\begin{align*}
abc & \quad abd & \quad acb & \quad acd & \quad adb & \quad adc \\
\text{bac} & \quad \text{bad} & \quad \text{bca} & \quad \text{bcd} & \quad \text{bda} & \quad \text{bdc} \\
\text{cab} & \quad \text{cad} & \quad \text{cba} & \quad \text{cbd} & \quad \text{cda} & \quad \text{cdb} \\
\text{dab} & \quad \text{dac} & \quad \text{dba} & \quad \text{dbc} & \quad \text{dca} & \quad \text{dcb}
\end{align*}
\]

Would you want to do this for 7 from a set of 10?
Combinations

Choosing subsets – order doesn’t matter.
How many ways can you do it?
Combinations of $k$ from a set of $n$

Give all combinations of 3 things out of \{a, b, c, d\}

Answer: \{a,b,c\}, \{a,b,d\}, \{a,c,d\}, \{b,c,d\}
Permutations and Combinations

\[\begin{align*}
abc & \quad acb & \quad bac & \quad bca & \quad cab & \quad cba & \{a, b, c\} \\
abd & \quad adb & \quad bad & \quad bda & \quad dab & \quad dba & \{a, b, d\} \\
acd & \quad adc & \quad cad & \quad cda & \quad dac & \quad dca & \{a, c, d\} \\
bcd & \quad bdc & \quad cbd & \quad cdb & \quad dbc & \quad dcb & \{b, c, d\}
\end{align*}\]

Permutations: \[4P_3\]

Combinations: \[\binom{4}{3} = 4C_3\]

\[\binom{4}{3} = 4C_3 = \frac{4P_3}{3!}\]
(a) Count the number of ways to get exactly 3 heads in 10 flips of a coin.

(b) For a fair coin, what is the probability of exactly 3 heads in 10 flips?

**answer:** (a) We have to 'choose' 3 out of 10 flips for heads: \[ \binom{10}{3} \].

(b) There are \( 2^{10} \) possible outcomes from 10 flips (this is the rule of product). For a fair coin each outcome is equally probable so the probability of exactly 3 heads is

\[
\frac{\binom{10}{3}}{2^{10}} = \frac{120}{1024} = .117
\]