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 2018

ALL SPORTS COMMENTARY

http://xkcd.com/904/

Staff

- David Vogan
  dav@math.mit.edu, office hours *Sunday 2–4* in *2-355*
- Nicholas Triantafillou
  ngtriant@mit.edu, office hours *Friday 5–7 p.m.* in *2-239A*
- Guangyi Yue
  gyyue@mit.edu, office hours *Sunday 8–10 a.m. room 2-333A*
- Richard Zhang
  zrichard@mit.edu, office hours *Saturday 2–4 room 2-490*

Web site

- 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.
- 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/board questions
  No computer use in class on MW.
- 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
- Problem sets will be graded on the logic and explanation of your answer.
- Just writing a number is not enough!

R, Piazza

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 web site.

Calendar, Information, Policies and Goals

Everything above and more is **on the web site!** rtfw!

For Next Time

- Familiarize yourself with the web site
- 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
- Seek probabilistic conclusions from experimental data
- No single correct answer

Counting: Motivating Examples

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*: $\heartsuit, \spadesuit, \diamondsuit, \clubsuit$,

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 ranks
- Example: $\{2\heartsuit, 2\spadesuit, 5\heartsuit, 8\clubsuit, K\diamondsuit\}$

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%

Board question: Poker Hands

- Wikipedia https:en.wikipedia.ord/wiki/Poker_probability informs us that
  1. There are 2,598,960 poker hands.
  2. There are 1,349,088 hands containing exactly two cards of some rank (so a pair, but not three of that kind).
  3. There are 123,552 hands containing two pairs (but not a full house).
  4. There are 3,744 hands containing a full house.
  5. There are 1,098,240 hands containing a pair and nothing more.
- The number in **5** is the number in **2**, after subtracting full-house-hands and two-pair-hands.
- $1349088 - 123552 - 3744 = 1221792 \neq 1098240$

What's wrong, Wikipedia?

What's (not) wrong with Wikipedia

1. There are 2,598,960 poker hands.
2. There are 1,349,088 hands containing exactly two cards of some rank (so a pair, but not three of that kind).
3. There are 123,552 hands containing two pairs (but not a full house).
4. There are 3,744 hands containing a full house.
5. There are 1,098,240 hands containg a pair and nothing more.

The hands in **2** prefer one pair (if there are two). So you need to subtract the two-pair hands with the *small* pair preferred, and then the two-pair hands with the *big* pair preferred:

$$1349088 - 123552 - 123552 - 3744 = 1098240.$$

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

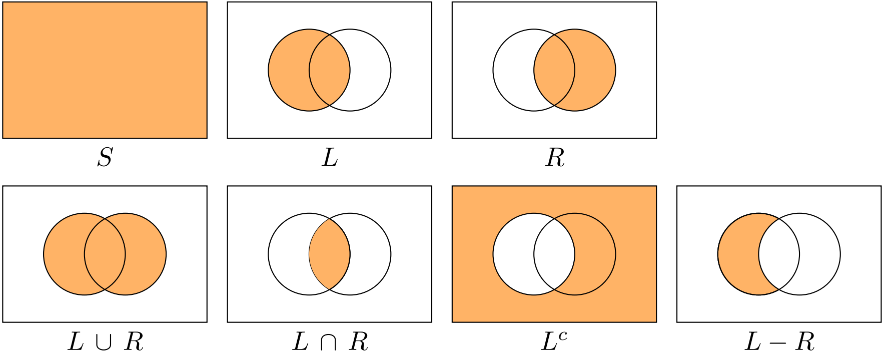
$S = \{$Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec$\}$

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

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

$L \cap R = \{$Jan, Mar, Oct, Dec$\}$ = months with 31 days and an 'r'

Visualize Set Operations with Venn Diagrams



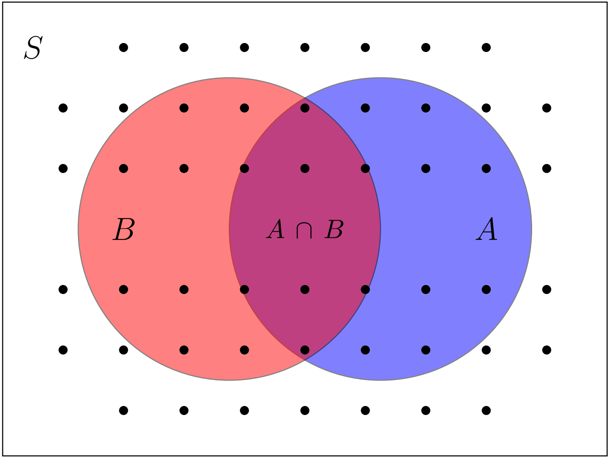
Product of Sets

$$S \times T = \{\text{pairs } (s, t) \text{ with } s \text{ in } S,\ t \text{ in } T\}$$

SIZE of $S \times T$ = (size of S) $\cdot$ (size of T)

$$|S \times T| = |S| \cdot |T|.$$

Inclusion-Exclusion Principle



Board Question

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

(set of shirts) $\times$ (set of pants) = set of outfits

$$|S| \cdot |T| = |S \times T|$$

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

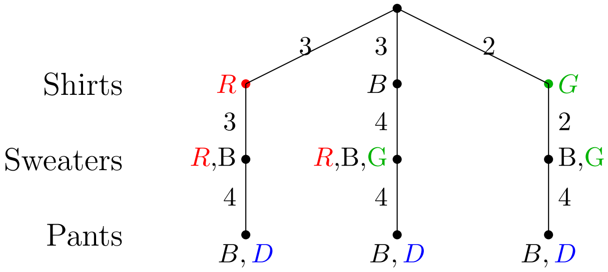
I won't wear green and red together; I think black or denim goes with anything; Here is my wardrobe.

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.



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' , 'cab' are 2 of the 6 permutations of {a, b, c}

'ad' , 'da' , 'bc' are three of the twelve permutations of two things from {a,b,c,d}

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 $\{a, b, c, d\}$

$$\begin{array}{cccccc} abc & abd & acb & acd & adb & adc \\ bac & bad & bca & bcd & bda & bdc \\ cab & cad & cba & cbd & cda & cdb \\ dab & dac & dba & dbc & dca & dcb \end{array}$$

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

$_nP_k$ =number of permutations (ordered lists)

of k things from n

$$_nC_k = \binom{n}{k}$$ =number of combinations (subsets)

of k things from n

$$_nP_k = \frac{n!}{(n-k)!} \qquad \binom{n}{k} = {}_nC_k = \frac{_nP_k}{k!}$$

Permutations and Combinations

abc	acb	bac	bca	cab	cba	$\{a, b, c\}$
abd	adb	bad	bda	dab	dba	$\{a, b, d\}$
acd	adc	cad	cda	dac	dca	$\{a, c, d\}$
bcd	bdc	cbd	cdb	dbc	dcb	$\{b, c, d\}$

Permutations:

$${}_4P_3$$

Combinations:

$$\binom{4}{3} = {}_4C_3$$

$$\binom{4}{3} = {}_4C_3 = \frac{{}_4P_3}{3!}$$

Board Question

(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: $\boxed{\binom{10}{3}.}$

(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

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