To the Members of the California State Assembly:

I am returning Assembly Bill 1176 without my signature.

For some time now I have lamented the fact that major issues are overlooked while many unnecessary bills come to me for consideration. Water reform, prison reform, and health care are major issues my Administration has brought to the table, but the Legislature just kicks the can down the alley.

Yet another legislative year has come and gone without the major reforms Californians overwhelmingly deserve. In light of this, and after careful consideration, I believe it is unnecessary to sign this measure at this time.

Sincerely, Arnold Schwarzenegger

Welcome to stat310! What does this message have to do with statistics?

Stat310

Welcome & intro to probability

Hadley Wickham

- 1. A message from the Governator
- 2. About me & course syllabus
- 3. Introduction to probability
- 4. Next time



To the Members of the California State Assembly:

I am returning Assembly Bill 1176 without my signature.

For some time now I have lamented the fact that major issues are overlooked while many unnecessary bills come to me for consideration. Water reform, prison reform, and health care are major issues my Administration has brought to the table, but the Legislature just kicks the can down the alley.

Yet another legislative year has come and gone without the major reforms Californians overwhelmingly deserve. In light of this, and after careful consideration, I believe it is unnecessary to sign this measure at this time.

Sincerely, Arnold Schwarzenegger To the Members of the California State Assembly:

am returning Assembly Bill 1176 without my signature.

For some time now I have lamented the fact that major issues are overlooked while many unnecessary bills come to me for consideration. Water reform, prison reform, and health care are major issues my Administration has brought to the table, but the Legislature just kicks the can down the alley.

Yet another legislative year has come and gone without the major reforms Californians overwhelmingly deserve. In light of this, and after careful consideration, I believe it is unnecessary to sign this measure at this time.

Sincerely, Arnold Schwarzenegger

http://gov.ca.gov/pdf/press/2009bills/AB1176_Ammiano_Veto_Message.pdf

A Democratic assemblyman who heckled the governor during a recent event in San Francisco actually received two messages: the veto letter itself and a not-so-subtle rebuke creatively hidden within it.

Like a find-the-word puzzle, the second message was visible by stringing together the first letter of each line down the left-hand margin. It consisted of a common four-letter vulgarity followed by the letters "y-o-u."

"My goodness. What a coincidence," said Schwarzenegger spokesman Aaron McLear. "I suppose when you do so many vetoes, something like this is bound to happen."

http://www.huffingtonpost.com/2009/10/27/schwarzenegger-sends-lawm_n_336319.html

A Democratic assemblyman who heckled the governor during a recent event in San Francisco actually received two messages: the veto letter itself and a not-so-subtle rebuke creatively hidden within it.

Like a find-the-word puzzle, the second message was visible by stringing together the first letter of each line down the left-hand margin. It consisted of a common four-letter vulgarity followed by the letters "y-o-u."

"My goodness. What a coincidence," said Schwarzenegger spokesman Aaron McLear. "I suppose when you do so many vetoes, something like this is bound to happen."

Is it?

http://www.huffingtonpost.com/2009/10/27/schwarzenegger-sends-lawm_n_336319.html

Your turn

What does "bound to happen" mean in this context? Do you agree that this acrostic could have happened by chance? How could you support your belief?

Turn to the person next to you and discuss for 2 minutes. I'll then ask a couple of you to share your thoughts.

Inspired by http://statistics.berkeley.edu/~stark/Preprints/acrosticVeto09.htm

Before we can answer these questions, we need to some new vocabulary and tools...

Before we can answer these questions, we need to some new vocabulary and tools...

But before we do that, I should introduce myself and give you some details about this class







I have two dogs



I like to bake

At Rice, I'm a major advisor for statistics, and a divisional advisor for McMurtry



In my free time, I play xbox games

Homework 0

Due Thursday 5pm

Make me a flash card so I can get to know you a little.

Photo on one side, name and 5 interesting facts about you on the other.

My goals

Provide you with a mathematical toolkit to reason about uncertainty.

Motivate why this is important and useful.

Provide an classroom environment designed to maximise the amount you learn.



If you only remember one thing, it should be this:

had.co.nz/stat310 (or just google stat310)

Textbook

Recommended, not required.

Mathematical Statistics with Applications, by Ramachandran and Tsokos

If you buy using amazon from this link, http://amzn.to/310-book, I get ~5% of the cost, which I'll put towards a class party. (If you buy anything from amazon, I'll get a cut)

10x homework + 3x stats in practice

Weekly homeworks let you practice what we covered in class. Three stats in practice essay help relate class to the real world

60% of grade. Lowest two dropped.

Due 4pm Thursday in stat310 mail box in DH 1091. 10% penalty if turned in before 5pm Friday. Not accepted after that. (No exceptions)



Honour code

You may co-operate on homework but you must submit your own assignment that reflects your own thinking, work and organisation.

To check if your homework meets this standard, imagine I asked you to explain your reasoning for each problem - you should be able to do so with ease.

All homeworks are considered to be pledged to this standard.

Tests + Final

Two take home tests (10% each), plus one cumulative in-class final (20%).

Will allow one page of notes for each test, three pages for final.



By Day 2, if you have done nothing with the information you learned in that lecture, didn't think about it again, read it again, etc. you will have lost 50%-80% of what you learned. Our brains are constantly recording information on a temporary basis: scraps of conversation heard on the sidewalk, what the person in front of you is wearing. Because the information isn't necessary, and it doesn't come up again, our brains dump it all off, along with what was learned in the lecture that you actually do want to hold on to!

By Day 7, we remember even less, and by Day 30, we retain about 2%-3% of the original hour! This nicely coincides with midterm exams, and may account for feeling as if you've never seen this before in your life when you're studying for exams - you may need to actually re-learn it from scratch.

http://www.adm.uwaterloo.ca/infocs/study/curve.html

Help sessions

From next week, will run evening homework help sessions on Tuesday and Wednesday.

NOT a repeat of the lecture. Informal.

Opportunity for one-on-one help with homework or other questions from class.

Grading

- A = 90 100
- B = 80 89
- C = 65 79
- D = 50 64

F if you get less than 50% in any assessment category (homeworks, tests and final).

+ and - will be awarded at my discretion.
A+ requires exceptional performance.

Your turn

You'll get opportunities to practice what we've just talked about in class with "your turns".

1-5 minutes of you working

I'll call on someone by name to answer the problem. This will be moderately stressful.



http://en.wikipedia.org/wiki/File:HebbianYerkesDodson.JPG

Topic	Mathematical tool
Probability	Set theory
Discrete r.v.	Sums
Continuous r.v.	Univariate calculus
Multivariate r.v.	Multivariate calculus
Statistical inference	All of the above

Math background

Set theory, algebra, calculus See quick reference sheet online Not guaranteed to be exhaustive

We will make extensive use of wolfram alpha.

Back to the problem

Did this message To the Members of the Califor arise by chance?

am returning Assembly Bill 1176 without my signature.

For some time now I have lamented the fact that major issues are overlooked while many unnecessary bills come to me for consideration. Water reform, prison reform, and health care are major issues my Administration has brought to the table, but the Legislature just kicks the can down the alley.

Yet another legislative year has come and gone without the major reforms Californians overwhelmingly deserve. In light of this, and after careful consideration, I believe it is unnecessary to sign this measure at this time.

Sincerely, Arnold Schwarzenegger

http://gov.ca.gov/pdf/press/2009bills/AB1176_Ammiano_Veto_Message.pdf

What is probability?

Probability is the mathematical machinery necessary to answer questions about **uncertain events**.

We're scientists/engineers so we need to make this precise...

Random experiment

"A random experiment is an experiment, trial, or observation that can be repeated numerous times under the same conditions... It must in no way be affected by any previous outcome and cannot be predicted with certainty." (http://cnx.org/content/m13470/latest/)

i.e. it is **uncertain** (we don't know ahead of time what the answer will be) and **repeatable** (ideally).

Sample space

The **sample space** is the **set** containing all possible **outcomes** from a random experiment. Often called S.

(In set theory this is usually called U, but it's the same thing)

Your turn

For each of the following random experiments write down the sample space (1 minute).

Next, pair up and compare your answers. Can you order the experiments by the number of outcomes? (2 minutes)

I'll then ask a few of you (selected at *random*) to share your answers.

What is S?

The number of dice rolls until you roll a six. Your final letter grade in this class. The length of time until you next sneeze. The result of a coin flip. The next President of the United States. The weight of a bag of m&m's.

Size

- Small (< 10)
- Large, but finite
- Countably infinite
- Uncountably infinite

We will follow this order as we develop increasingly complex mathematical tools

Events

An event is a subset of the sample space. Events are usually given capital letters. The probability of an event A is written P(A).

More terminology

A collection of events are **mutually exclusive** if...

A collection of events are exhaustive if...

A collection of events is a **partition** if it is mutually exclusive and exhaustive

l f**k you

If we want to frame this as a probability problem, what might the experiment be? What is the sample space? What partition of the sample space might we be interested in?

One answer

- Sample space: all possible letters written by Arnold Schwarzenegger
- Partition of interest: presence/absence of offensive message

Probability

If I flip a fair coin, the probability of it coming up heads is 50% (or 0.5).

$$S = \{H, T\}$$
. $P(H) = 0.5$

But what does that mean?

Defining probability

- Informally
- Classically
- Frequentist
- **Axiomatic** (i.e. we don't care what it means, we just want it to behave in a sensible way)



Let A, B, C a partition of S. Which of the following don't make sense? P(A) > P(B)P(A) + P(B) + P(C) > 1P(S) = 0.9P(A) + P(B) = P(C) $P(A \cup B) = P(A) + P(B) + 0.1$ P(A) = -0.1

Three axioms:

Probability is a function, P, that satisfies these conditions:

- 1. $P(A) \ge 0$, for all $A \subset S$
- 2. P(S) = 1
- 3. $P(A \cup B) = P(A) + P(B)$ if $A \cap B = 0$

Any function that satisfies these axioms is called a probability function.

Next time

We'll start with a very simple class of events: those with equally likely outcomes. We'll look different probability models (approximations) and see if that message is likely or not.