Progress Report 1

Side note

I missed the first class because I wasn't sure whether I'd end up taking Bayesian or not. Now I'm sure that I'll take the class, but I still feel a little behind, and I think this progress report reflects that. I haven't had time to really sit down with `thinkbayes2.py` yet and understand it. And because I'll be gone from Friday-Sunday for a hackathon, I'm going to miss a bunch of time (in addition to class on Friday) that I would have spent catching up.

I have been able to keep pretty much caught up on the reading, and I feel like I have a good grasp of the fundamental concepts right now. My schedule doesn't let me work in large blocks until Wednesday/Thursday, so I can't say I'll be fully caught up by Tuesday, but I hope to be back on track by Friday. I think the material is cool and I want to put in the time to learn it.

1. Reading

I read the first three chapters of *Think Bayes* and skimmed the recommended Yudkowsky reading. I liked the chapters of *Think Bayes* -- I thought they were concise and easy to understand. I think it would have been cool to delve a bit further into the German tank problem. I liked the practice problems from the Yudkowsky article, but I found the article itself extremely long and repetitive. If I had read every word it would have taken me hours and added little additional value. I also read the *Of Miracles* wikipedia link, which pleased me.

2. Exercises

I did some dice exercises:

```
In [4]: dice = tb.Pmf()
In [5]: for i in range(1, 7):
   ...:     dice.Set(i, 1)
   ...:
In [6]: dice.Normalize()
In [7]: tp.Hist(dice)
In [8]: dice2 = dice + dice
In [9]: dice3 = dice2 + dice
In [10]: dice4 = dice3 + dice
In [11]: dice5 = dice4 + dice
In [13]: tp.Hist(dice2)
In [14]: tp.Hist(dice3)
In [15]: tp.Hist(dice4)
In [16]: tp.Hist(dice5)
In [17]: dice3.Var()
Out[17]: 8.749999999999998
In [18]: dice3.Mean()
Out[18]: 10.500000000000002
In [19]: tb.Std(dice3)
Out[19]: 4.6097722286464435
```
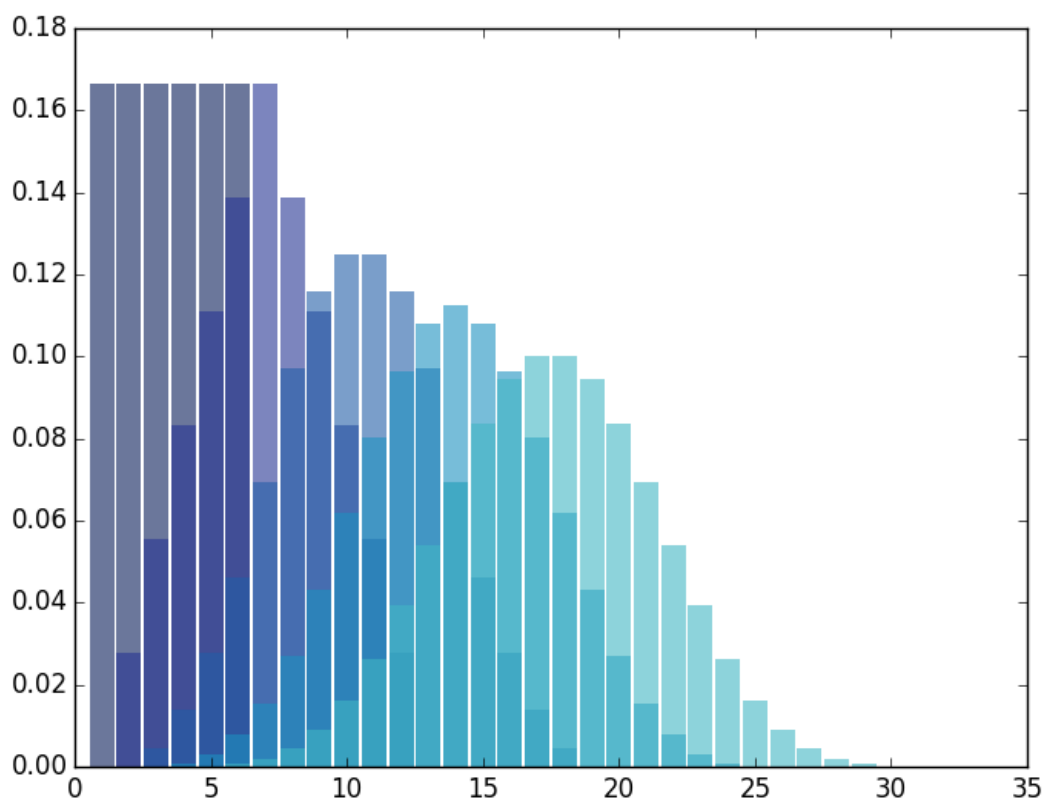
Plotting all the dice showed how they approach a normal distribution, which was pretty cool.

I also did quite a few exercises by hand -- this was while I was getting my environment set up. And either way, I find it faster and more intuitive to use the tree method over IPython for simple problems (like the egg/pearl one). I'm sure I'll begin to favor IPython as the problems get more complex.
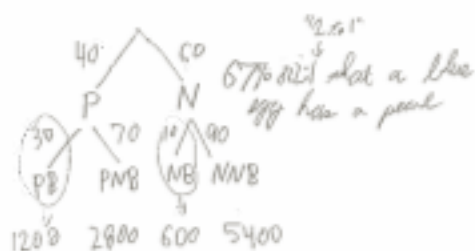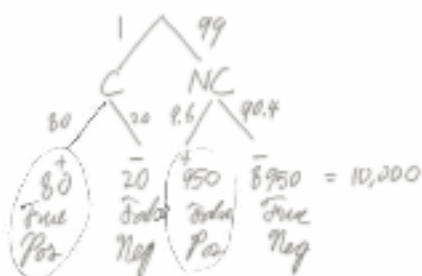
9/8/14

Prior P(H)	Likelihood P(D|H)	P(H)P(D|H)	P(H|D)
A	1	80	80	7.8%
B	99	9.6	950.4	92.2%

$$P(H|D) = \frac{P(H)\,P(D|H)}{P(D)}$$

P(D) = 1030.4

Tree method

1 /\ 99
C   NC
80 / \ 20   9.6 / \ 90.4

(80⁺) True Pos   20 False Neg   (950 Total Pos)   950 True Neg   = 10,000

"2.61"

40 / \ 60
P   N
30 / \ 70   10 / \ 90
(PB) PNB   (NB) NNB

67% 812.1 that a blue egg has a pearl

1200   2800   600   5400

Cookie Problem

$$P(H|D) = \frac{P(H)\,P(D|H)}{P(D)}$$

30V 10C   20V 20C

1   2

random bowl, random cookie ⇒ $\underset{\text{data}}{\underline{\text{vanilla}!}}$

$D = \text{"} \textit{2 chose vanilla"} = \frac{50}{80} = 5/8$

$\underset{P(D)}{}$

H = I think $\underline{\text{vanilla}}$ come from $\underline{\text{bowl 1}}$

$P(H) = \text{"}$ chance that it came from bowl $1 = 1/2$

$P(D|H) = $ probability that I chose vanilla given that I chose from bowl 1

$= \frac{30}{40} = 3/4$

$P(H|D) = $ prob that I chose from bowl 1 given that I chose vanilla

$= \frac{(1/2)(3/4)}{5/8} = \frac{3/8}{5/8} = \boxed{\dfrac{3}{5}}$

3. Case Study

I skimmed the cancer case study reading, and I remember the Angelina Jolie case being mentioned in class. I don't think we've done much with the case study yet, so this is about it.

4. Reflection

I got a much better sense for the meaning of Bayes' theorem from doing the cookie exercise by hand. I missed the first class... did we make learning goals? If I had to name a couple of learning goals, they might be:

- Gain the ability to use Bayesian reasoning to improve my understanding of the world and world events (and the things the media reports, etc.)
- Improve my ability to use Python as a tool to solve one-off problems (especially with IPython)-- basically wean myself off of MATLAB