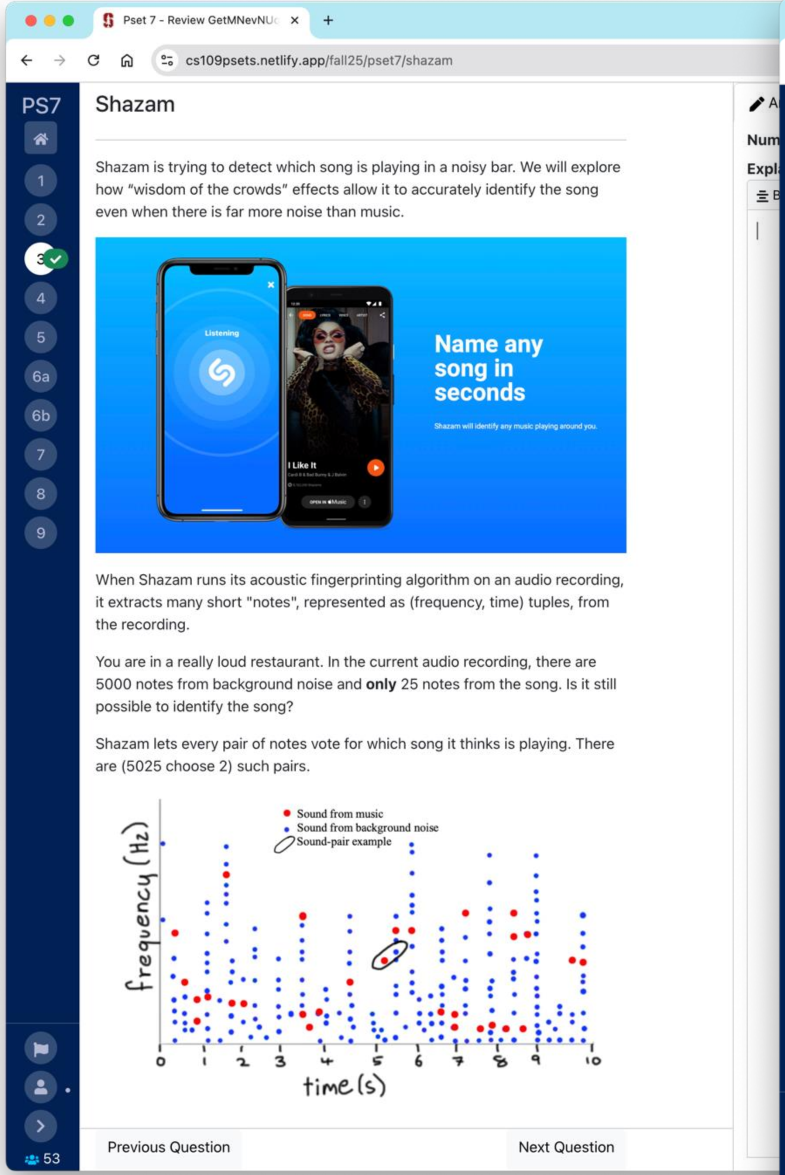
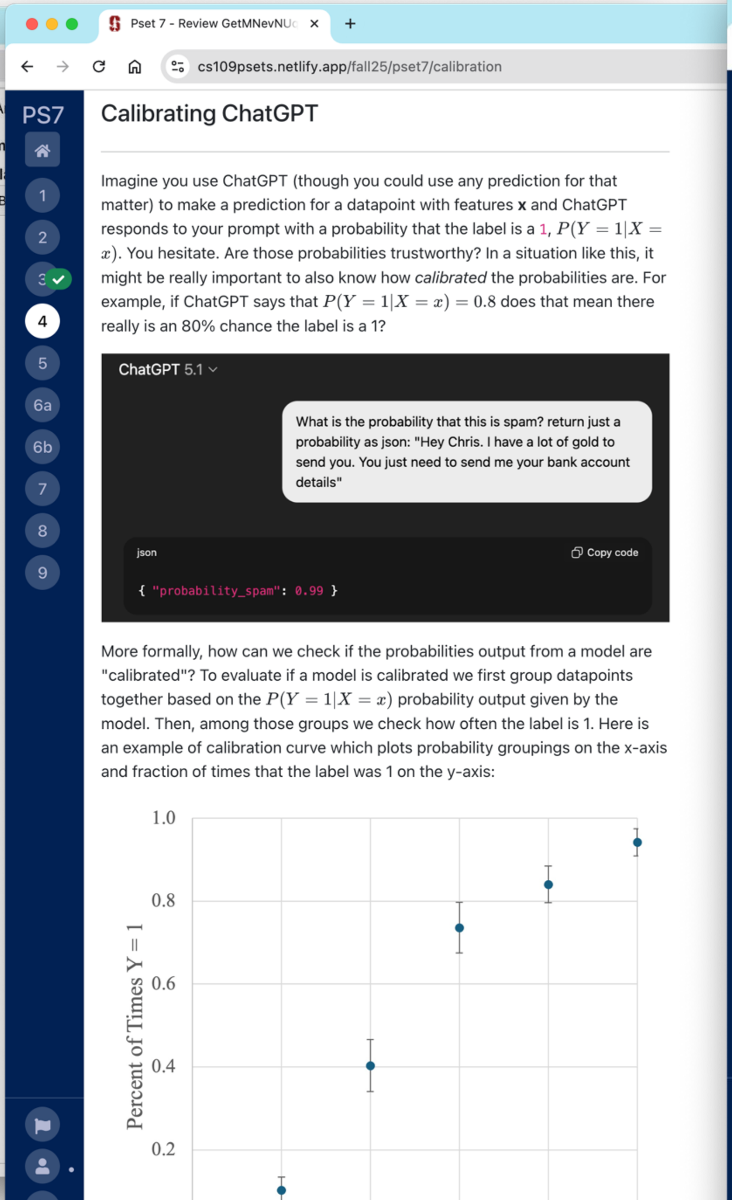
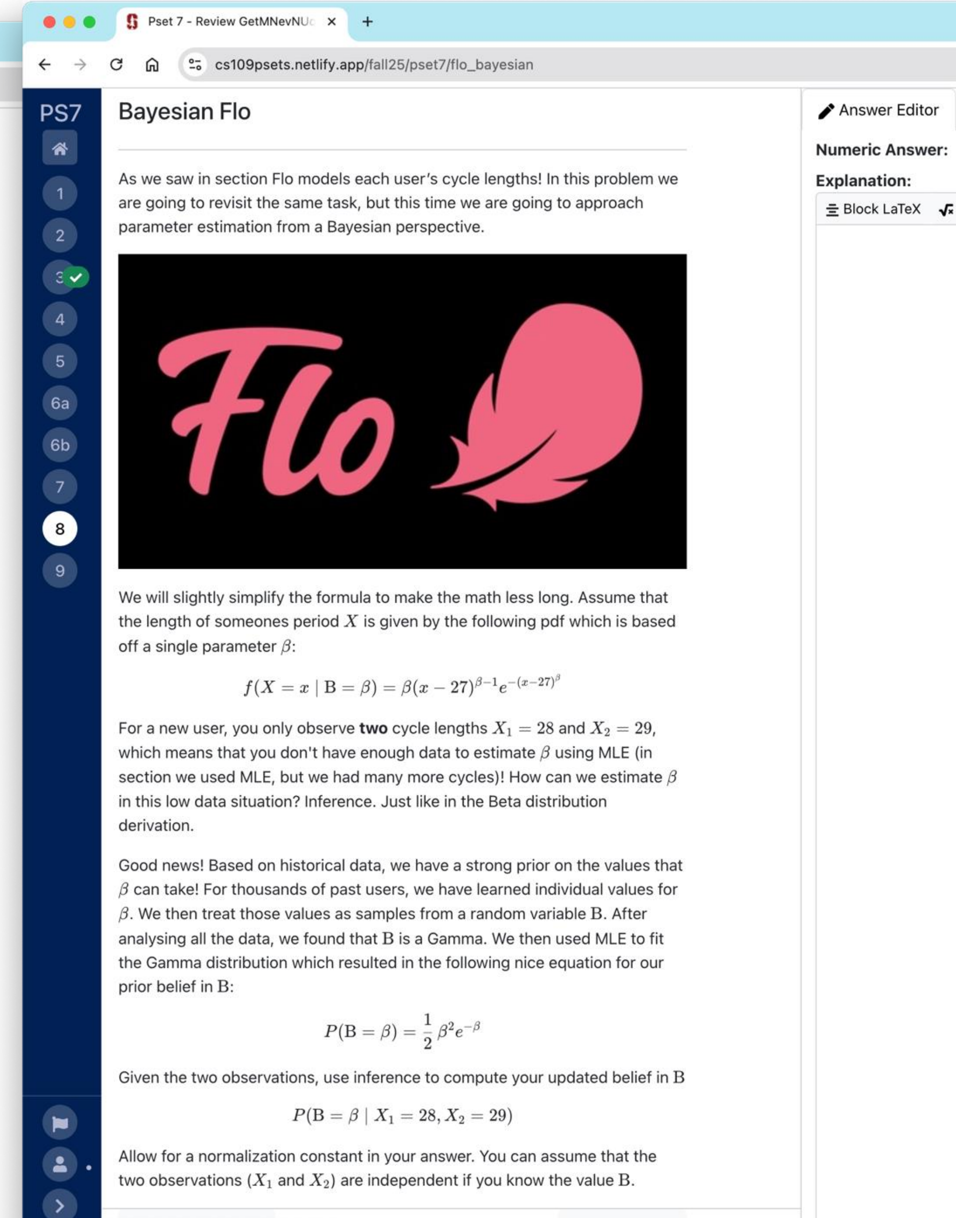
Reinforcement 2

Chris Piech
CS109, Stanford University

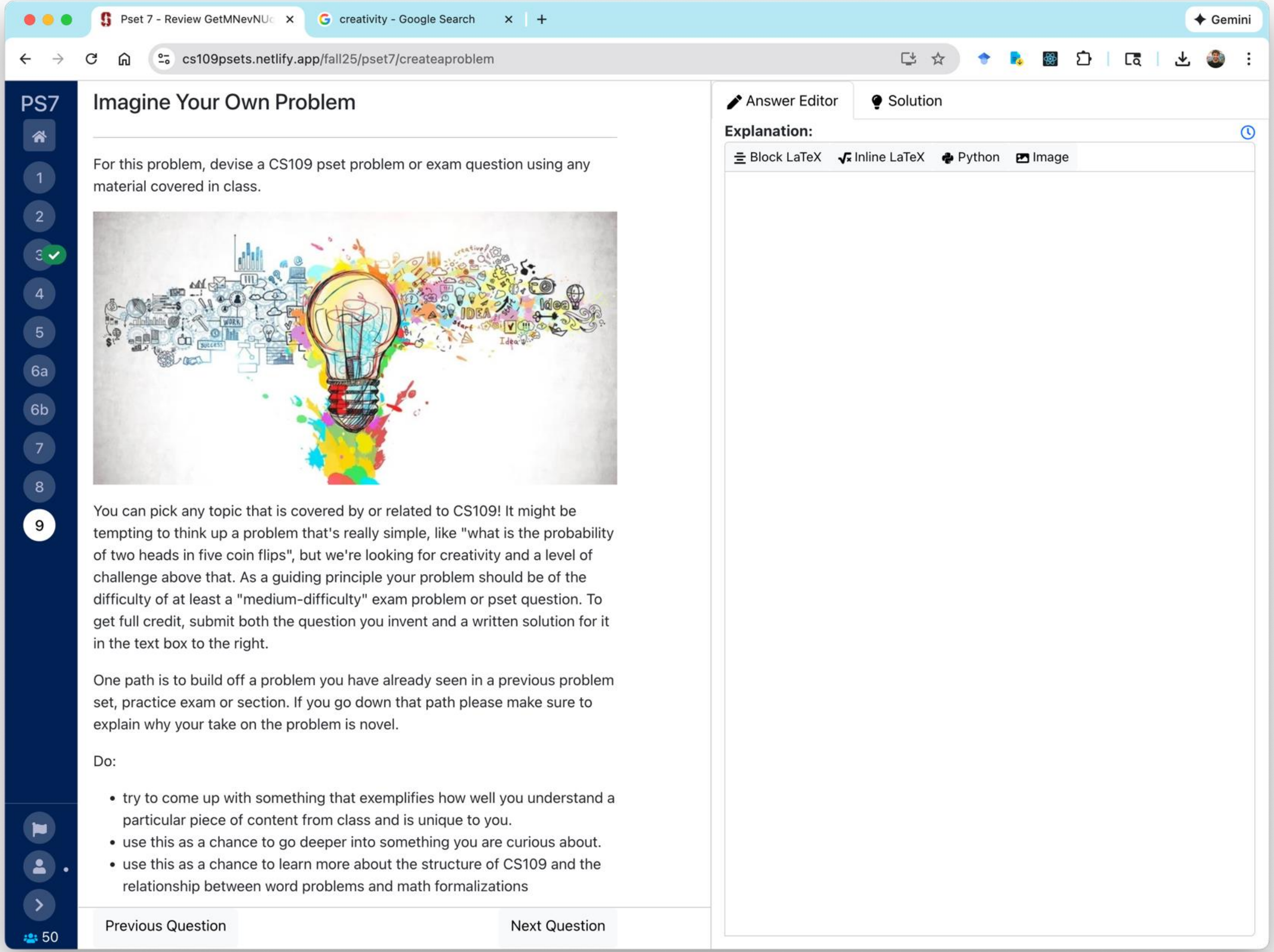
Pset 7: Review







Pset 7: Review



Today

What to Buy?

Poker Information Theory

Quant Interview (with cupcakes)

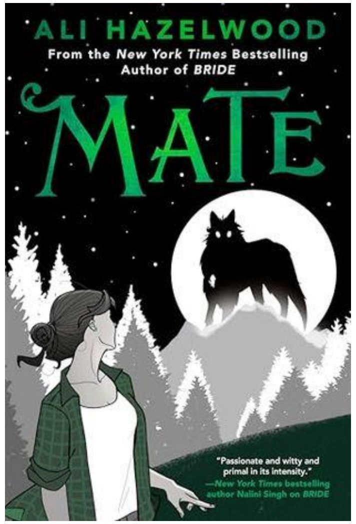
Child word identification

Summarize Medical Record

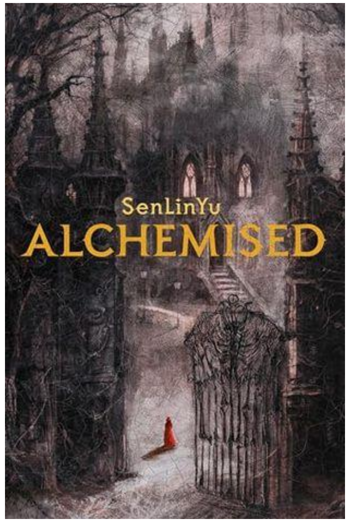
Beta Binomial

Rating Preference

Which Book Would you Prefer?



Rating	1	2	3	4	5
Count	0	0	0	1	4



Rating	1	2	3	4	5
Count	0	1	2	20	200

MLE Estimate

Which Book Would you Prefer? MLE Solution

$$P(R_2 > R_1) = \sum_{i=1}^{5} \sum_{j=i+1}^{5} P_1(i)\, P_2($$

```python
import numpy as np

# MLE probabilities for each book
# Order: ratings 1, 2, 3, 4, 5

p1 = np.array([0, 0, 0, 1/5, 4/5])             # Book 1
p2 = np.array([0, 1/223, 2/223, 20/223, 200/223])  # Book 2

prob = 0.0
for i in range(5):           # rating for Book 1 (0-index → rating i+1)
    for j in range(i+1, 5):   # rating for Book 2 strictly greater
        prob += p1[i] * p2[j]

print("P(Book 2 > Book 1) =", prob)
```

Another Perspective: Beta, but Better

Which Book Would you Prefer? Better Beta Solution

Assume: Rating probabilities are distributed as a Dirchlet with the following joint PDF:

$$f(x_1, \ldots, x_5) = K \cdot \prod_{i=1}^{5} x_i^{c_i}$$

How many times you saw rating i

Probability that you will get rating 1

Probability that you will get rating k

You can assume you have access to

```
# take a sample from the joint
x_list = sample_dirchlet()

# evaluate the joint pdf
likelihood = pdf(x_list)
```

Another Perspective: Latent Like

Book vibe

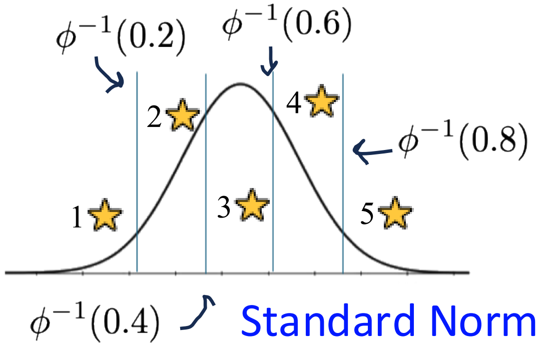
$$Y \sim N(\mu = 0.8, \sigma^2 = 1)$$

$$X_i \sim N(\mu = Y, \sigma^2 = 1)$$

Rater i vibe of the book

Could fit these using historical data!

Rating	1	2	3	4	5
Count	0	0	0	1	4



Standard Normal

$$P(\text{Ratings}|Y = y) = P(4\star|Y = y) \cdot P(5\star|Y = y)^4$$

Threshold 3 Threshold 4

$$P(\tau_3 < X_i < \tau_4 | Y = y)$$

$$= P\left(\frac{\tau_3 - y}{1} < \frac{X_i - y}{1} < \frac{\tau_4 - y}{1}\right)$$

$$= P(\tau_3 - y < Z < \tau_4 - y)$$

$$= \Phi(\tau_4 - y) - \Phi(\tau_3 - y)$$

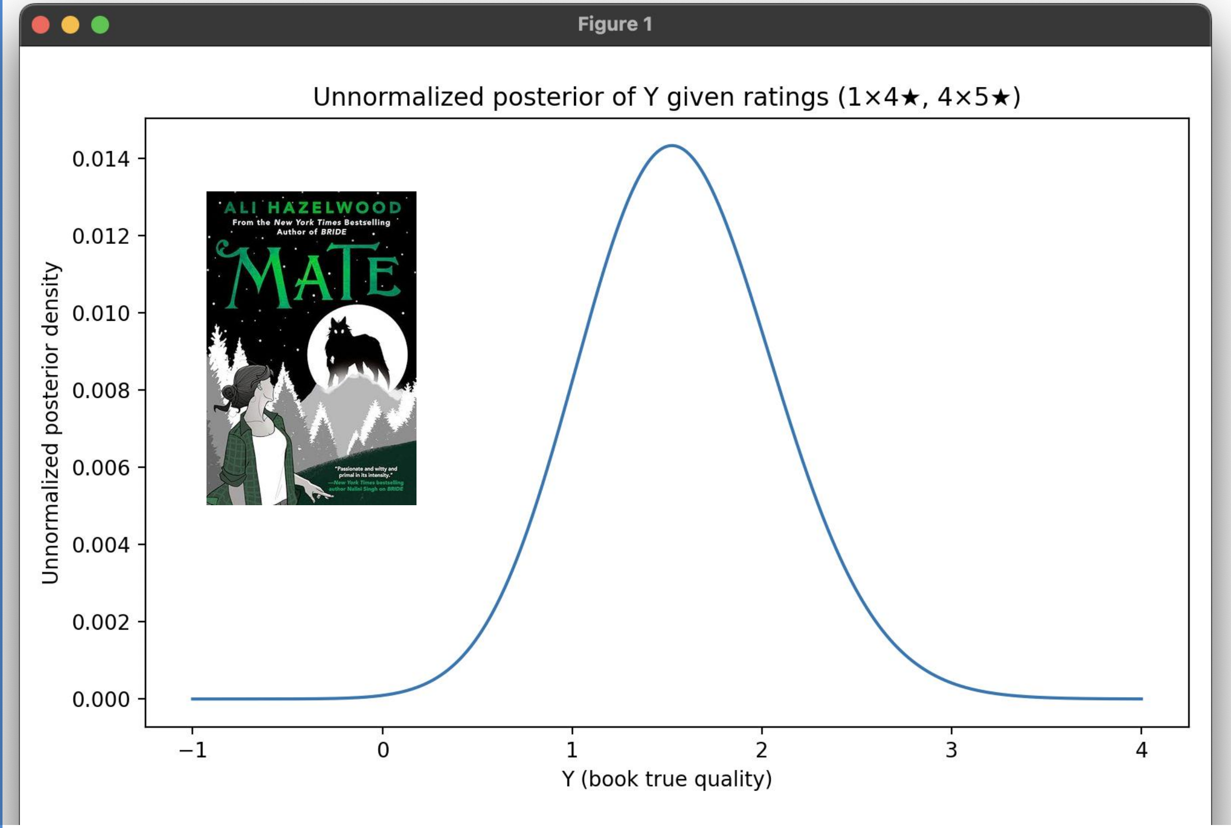
Normal PDF

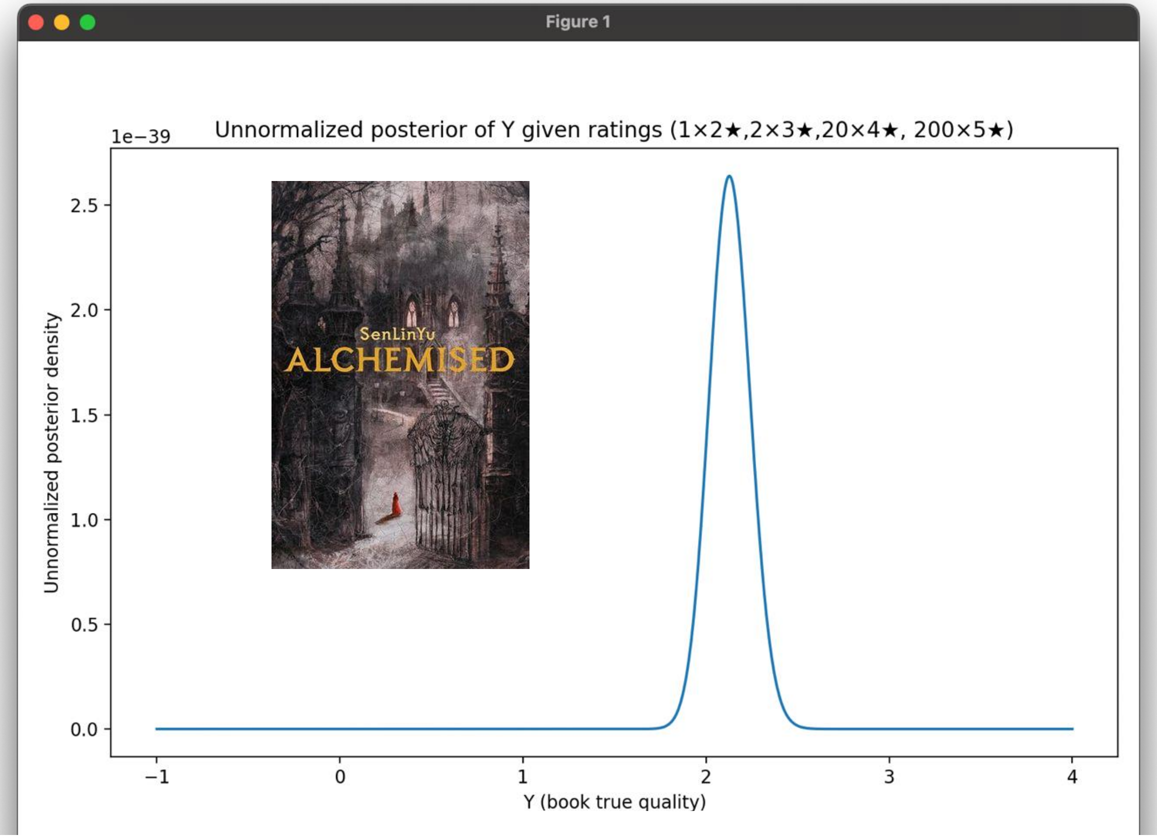
$$f(Y = y \mid \text{Ratings}) \propto \exp\left(-\frac{1}{2}(y - 0.8)^2\right) \cdot [\Phi(\tau_4 - y) - \Phi(\tau_3 - y)] \cdot [1 - \Phi(\tau_4 - y)]^4$$

$\Phi^{-1}(0.8)$ $\Phi^{-1}(0.6)$ $\Phi^{-1}(0.8)$

Which Book Would you Prefer?

Probability you prefer book 2 is 86%

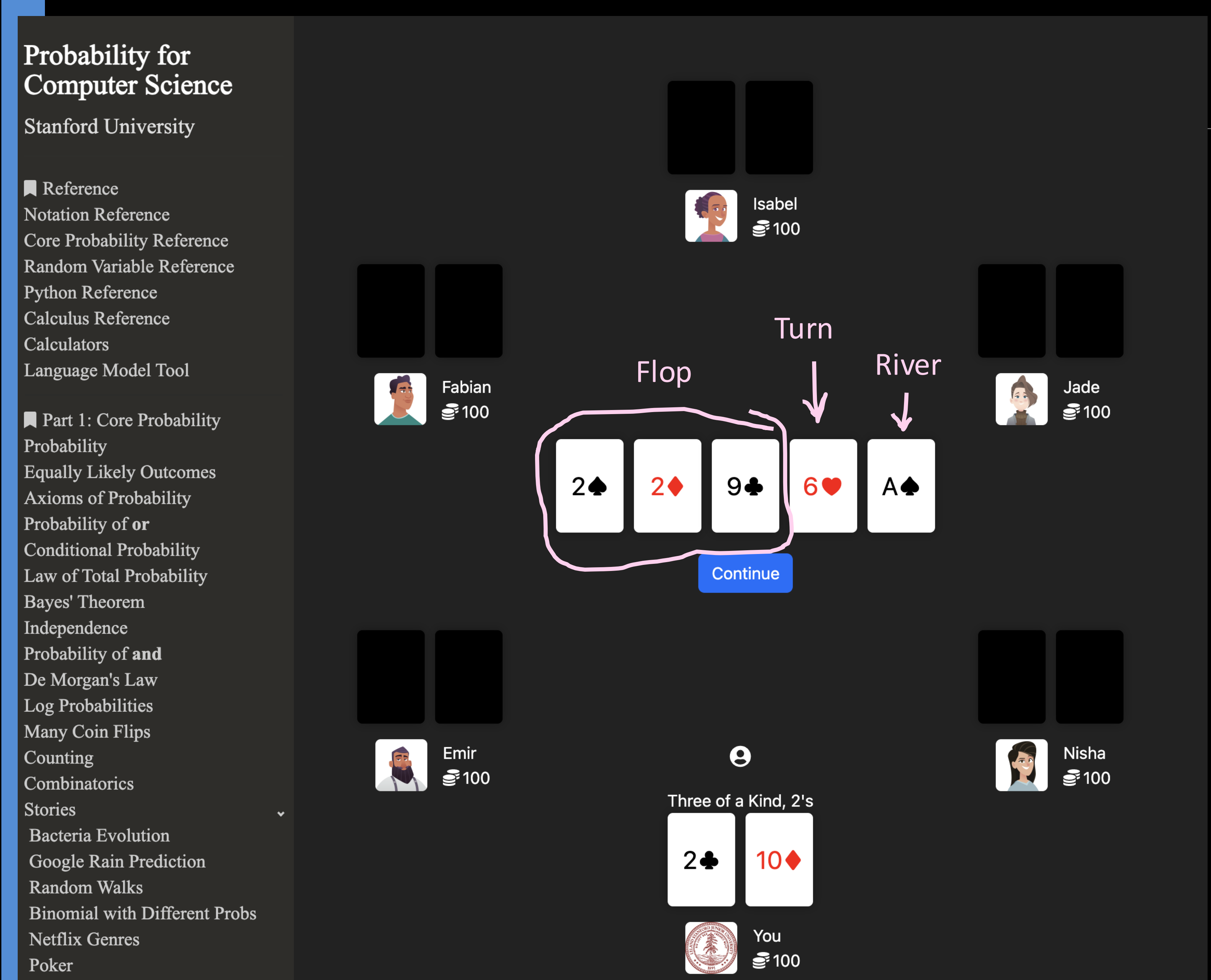




Rating	1	2	3	4	5
Count	0	0	0	1	4

Rating	1	2	3	4	5
Count	0	1	2	20	200

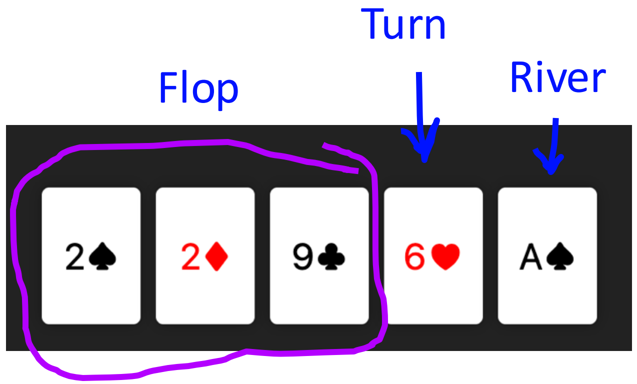
Poker Information Theory



Let X be a random variable for the set of 7 cards at the end. What is H(X) after each event?

There are 52 unique cards. Each card is equally likely.

Poker Entropy



H(pre flop) = 21.014

H(pre turn) = 10.078

H(pre river) = 5.524

H(post river) = 0

Quant Interview

Quant Interview

100 cupcakes.
Scenario 1: 51 are blue.
Scenario 2: 49 are blue.

a) You look at one cupcake. It is blue.

b) You look at three cupcakes (with replacement). Two are blue.

c) You look at three cupcakes (without replacement). Two are blue.

Quant Interview: Bonus

c. A hypergeometric is a random variable for the number of successes if you remove items **without** replacement from a fixed population. If $X \sim \text{Hypergeom}(t, k, n)$, the PMF is

$$P(X = x) = \frac{\binom{k}{x}\binom{t-k}{n-x}}{\binom{t}{n}}$$

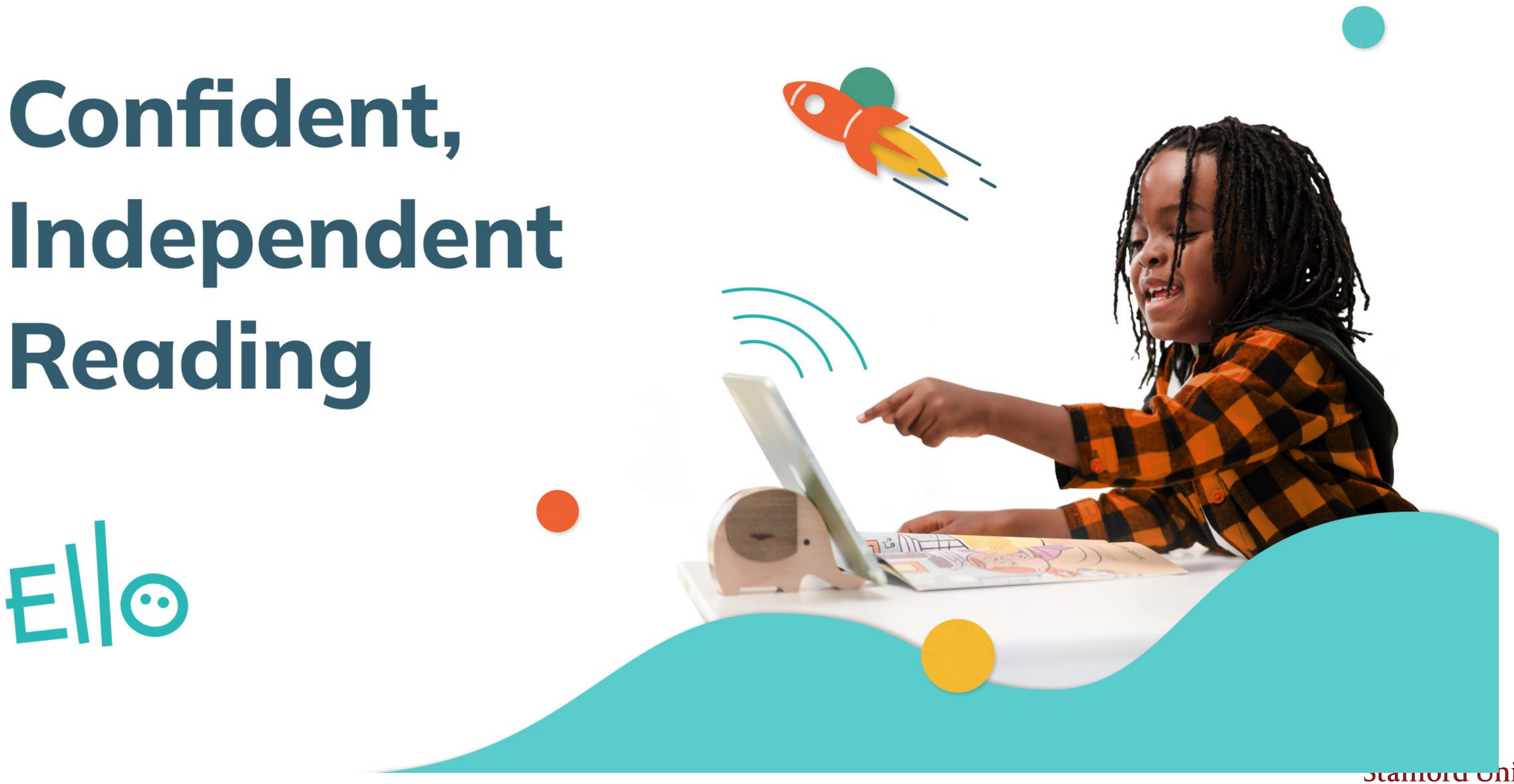
Where:

- t is the population size,
- k is the number of "success" items in the population,
- n is the number of draws (without replacement),
- X counts the number of successes observed in those n draws.

Three cupcakes are drawn without replacement and two are blue. What is the probability that the majority were blue?

Did they read that?

Did they read that?



Did they read that?

Medical Summary

2. What an Ideal Classifier Looks Like

We compute the true conditional probability $P(I = i \mid x, y_1, \ldots, y_N)$. Using conditional probability,

$$P(I = i \mid x, y_1, \ldots, y_N) = \frac{P(x, y_1, \ldots, y_N, i)}{P(x, y_1, \ldots, y_N)}.$$

Because the distractors Y_j for $j \neq i$ are drawn independently from $P(Y)$ and independently of X,

$$P(x, y_1, \ldots, y_N, i) = \frac{1}{N} P(x, y_i) \prod_{j \neq i} P(y_j).$$

The denominator is a sum over all possible positions of the true record:

$$P(x, y_1, \ldots, y_N) = \sum_{k=1}^{N} \frac{1}{N} P(x, y_k) \prod_{j \neq k} P(y_j).$$

How does this follow →

$$P(I = i \mid x, y_1, \ldots, y_N) = \frac{P(x, y_i)/P(y_i)}{\sum_{k=1}^{N} P(x, y_k)/P(y_k)}$$

$$= \frac{P(y_i \mid x)P(y_i)}{\sum_{k=1}^{N} P(y_k \mid x)/P(y_k)}$$