Naïve Bayesians

Back to Basics Series

16 Jan 2021

Goal

Developing the Bayesian muscle to solve a wide range of problems

Naïve Bayesian Philosophy

Intuitive (Visual) Understanding of the Bayesian Reasoning

Ability to model real world problems in a Bayesian Setting

Starting from Simple Probabilistic modelling

Adapting it in a a Bayesian setting
And moving towards ML models

Fluency in the Calculus of Bayesian Stats & ML model

Season 2: Back to Basics

Ep 1	Bayes Theorem
Ep 2, Ep 3	Problems with Binomial Likelihoods
Ep 4	Disease Detection
Ep 5	Naive Bayes Classification
Ep 6	Gaussian Naive Bayes Classification
Ep 7	German Tank Problem
Ep 8	Waiting Times (Continuous Distributions)

Back to Basics

		Canonical Problem	Applications
Ep 1	Bayes Theorem	There are 2 boxes from which cookies can be taken from. Box A and Box B. Box A contains 10 chocolate cookies, Box B contains 5 ginger cookies. Given that you get a chocolate cookie which box was it taken from?	The Shy Librarian Problem Naive Bayes algorithm
Ep 2 Ep 3	Problems with Binomial Likelihoods	You have 2 coins C1 and C2. p(heads for C1) = .7 & P(heads for C2) = 0.6 You flip the coin 10 times. What is the probability that the given coin you picked is C1 given you have 7 heads and 3 tails?	A/B Testing
Ep 4	Disease Detection	A particular disease affects 1% of the population. There is an imperfect test for this disease: The test gives a positive result for 90% of people who have the disease, and 5% of the people who are disease-free. Given a positive test result – what is the probability of having the disease?	COVID Tests (PCR & Antibody)! Fraud Detection
Ep 5	Naive Bayes Classification	Given these words occur in this text what's the probability it's spam?	Any Classification Problem
Ep 6	Gaussian Naive Bayes Classification	Given the weights and heights of basketball players, what's the probability that person a is a basketball player given weight = w and height = h?	

Back to Basics

		Canonical Problem	Applications
Ep 7	German Tank Problem	Suppose tanks were given a serial number based on the order in which they were manufactured. Given that you've observed a tank with serial number "10", how many tanks were actually manufactured in total?	?
Ep 8	Waiting Times (Continuous Distributions)	Suppose you need to gather 10 patients for a trial. Each signup happens at time t_i (i=1, 10). How long do you have to wait after it took you 3 weeks to accrue 2 signups?	Planning Trials Estimating Queues

Bayes Rule

Posterior Likelihood Prior

$$P(\theta \mid D) = \frac{P(D \mid \theta)\, P(\theta)}{P(D)}$$

Normalising Constant

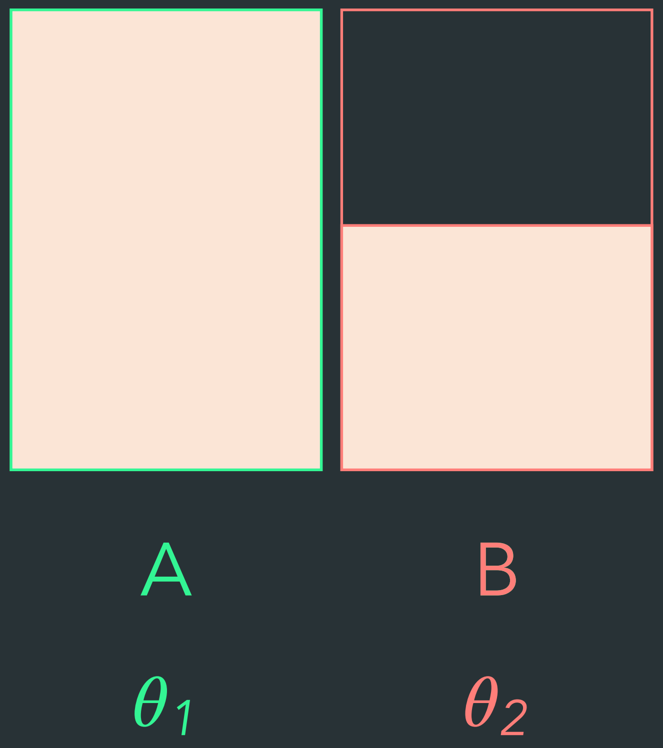
Bayes Rule

Posterior Likelihood Prior

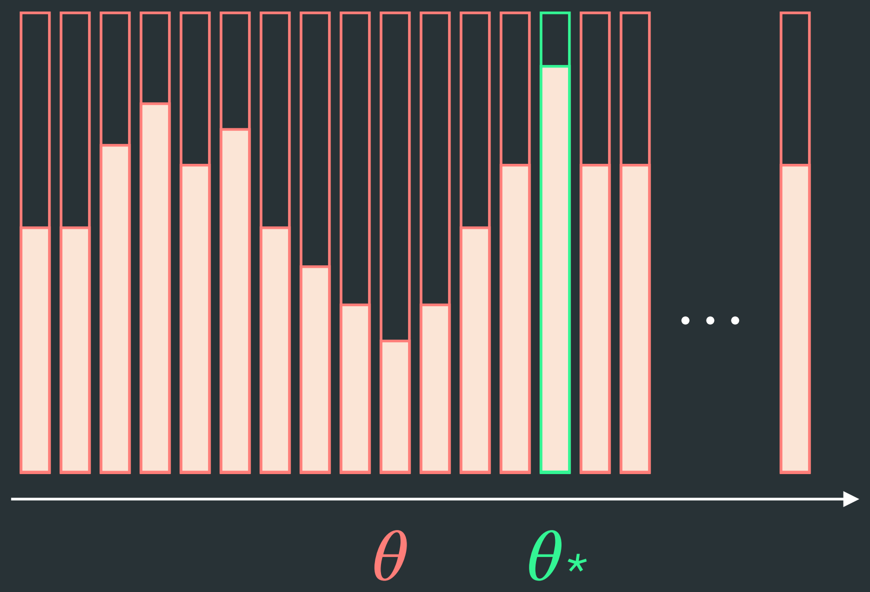
$$P(\theta_i \mid D) = \frac{P(D \mid \theta_i)\, P(\theta_i)}{\sum_{all\ j} P(D \mid \theta_j)\, P(\theta_j)}$$

Normalising Constant

Visual representation of Bayes Rule for 2 Hypotheses



What about continuous parameters?



Moving from a discrete set of hypothesis to a continuous parameters can be thought of having an infinite number of “boxes” to choose from

Bayes Rule

Posterior Likelihood Prior

$$P(\theta \mid D) = \frac{P(D \mid \theta)\, P(\theta)}{\int P(D \mid \theta')\, P(\theta')\, d\theta'}$$

Normalising Constant

Recap from Last Week | Canonical Problem

You have 2 coins C_1 and C_2.

$p(\text{heads for } C_1) = 0.7$ $p(\text{heads for } C_2) = 0.6$

You flip one of the coins 10 times and get **7 heads** and **3 tails**

What is the probability that the given coin you picked is C_1?

Canonical Problem

You flip one of the coins 10 times and get **7 heads** and **3 tails**

What is the probability of getting a head in the next flip?

Real world Problems to Solve

- Comparing Amazon Reviews
  - Seller A has 7 positive ratings with 10 reviews
  - Seller B has 70 positive ratings with 100 reviews
  - Seller C has 140 positive ratings with 200 reviews

- A/B Testing:
  - Webpage A has 30% conversion rate (1200 users)
  - Webpage B has a 35% conversion rate (1300 users)

7 out of 10 reviews are good what is the probability of getting a good review?

$n = 10, k = 7$

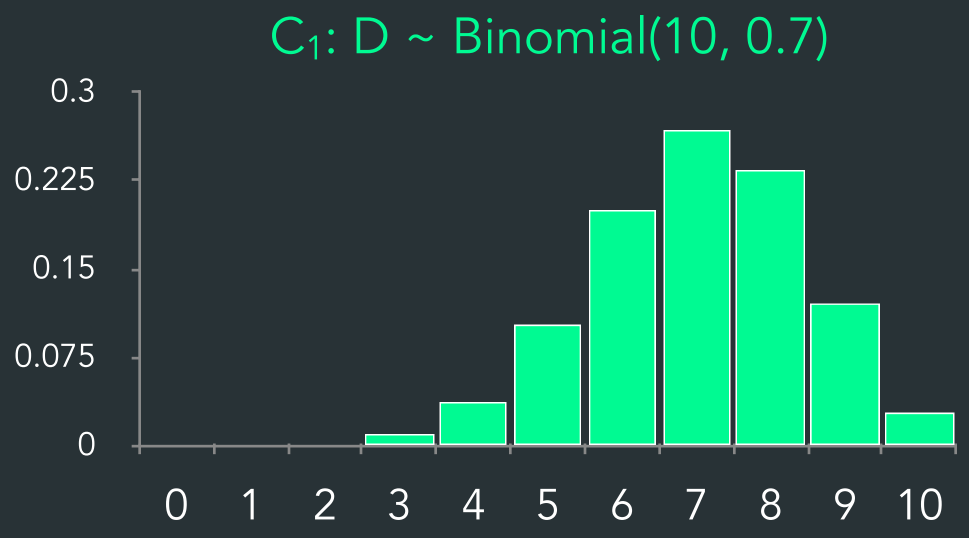
Expected value

- $7/10 = 70\%$

- How confident are you?
- What if 70 out of 100 reviews were positive?
- What if you knew that it was more likely for the data to tampered?

What is the probability that the given coin you picked is C_1?

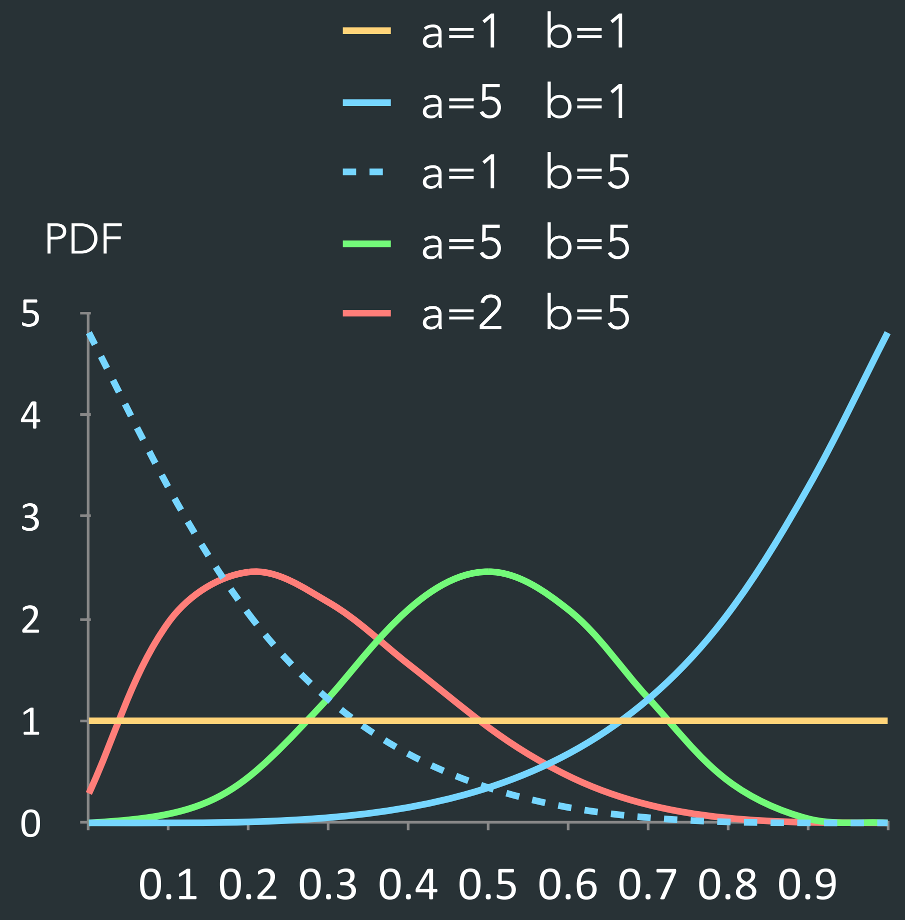
- D ~ Binomial(n, p)
  - $0 \leq D \leq n$
  - $n > 0$
  - $0 < p < 1$



$$P(D = k \mid p) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Beta Distribution

- $\theta \sim$ Beta(a, b),
  - $0 \leq \theta \leq 1$
  - $a > 0, b > 0$
- Flexible family of distribution
- Conjugate prior to the Binomial



$$P(\theta) = c\, \theta^{a-1} (1 - \theta)^{b-1}$$

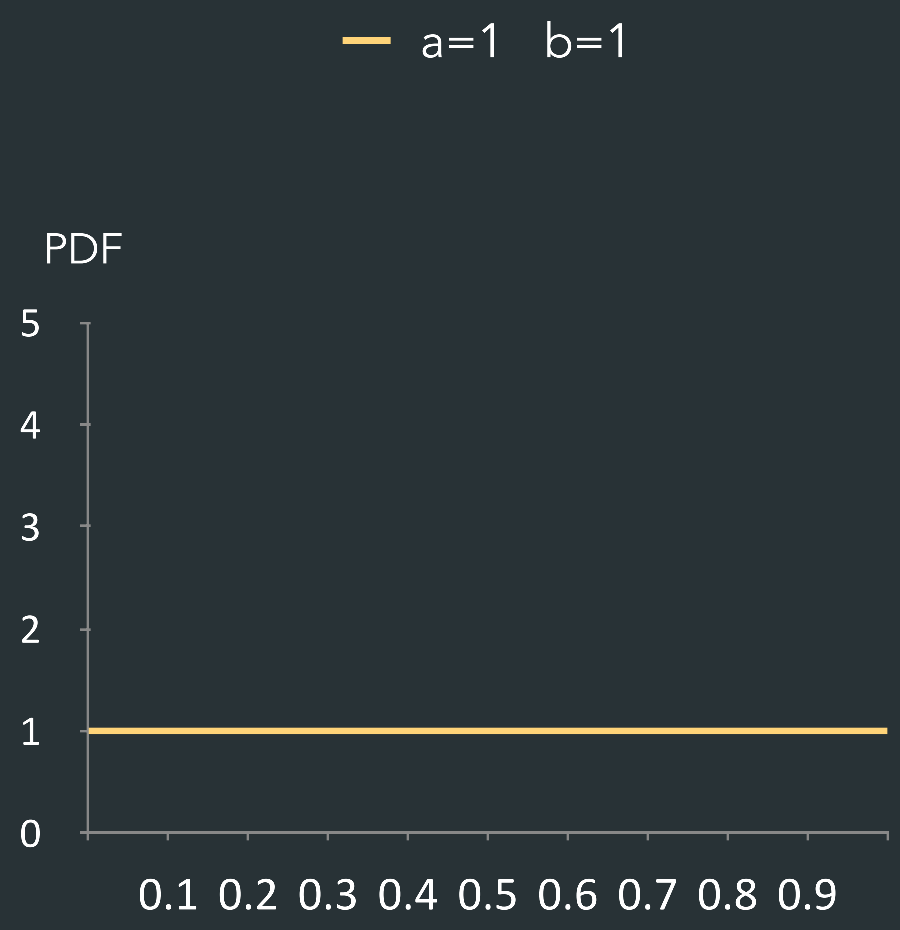
$$E(\theta) = \frac{a}{a + b} \qquad \text{Mode} = \frac{a - 1}{a + b - 2}$$

Beta Distribution

- $\theta \sim$ Beta(a, b),
  - $0 \leq \theta \leq 1$
  - $a > 0, b > 0$
- Flexible family of distribution
- Conjugate prior to the Binomial

$$P(\theta) = c\,\theta^{a-1}(1-\theta)^{b-1}$$

$$E(\theta) = \frac{a}{a+b} \qquad \text{Mode} = \frac{a-1}{a+b-2}$$



Bayes Rule

$$\underbrace{P(\theta \mid D)}_{\text{Posterior}} = \frac{\overbrace{P(D \mid \theta)}^{\text{Likelihood}}\ \overbrace{P(\theta)}^{\text{Prior}}}{\underbrace{P(D)}_{\text{Normalising Constant}}}$$

Conjugate Prior

Same Family

Posterior

Prior

$$P(\theta \mid D) = \frac{P(D \mid \theta)\, P(\theta)}{P(D)}$$

What is the probability of getting a head?

1. Let p be probability of getting a head
   - Assume p follows a Beta(1, 1) distribution

2. Let D be the number of heads in a set of 10 flips
   - Assume D follows a Binomial(10, p) distribution

$$P(p \mid D) = \frac{\overbrace{P(D \mid p)}^{\text{Binomial}(10,\ p)}\ \overbrace{P(p)}^{\text{Beta}(1,\ 1)}}{P(D)}$$

Bayes Rule

Posterior Likelihood Prior

$$P(\theta \mid D) = \frac{P(D \mid \theta)\, P(\theta)}{\int P(D \mid \theta')\, P(\theta')\, d\theta'}$$

Normalising Constant

What is the probability of getting a head?

$$P(\,p \mid D{=}k\,) \propto P(\,D{=}k \mid p\,)\,P(p)$$

$$\propto p^{k}(1-p)^{n-k}\,p^{a-1}(1-p)^{b-1}$$

$$\propto p^{a+k-1}(1-p)^{b+n-k-1}$$

$$= p^{a'-1}(1-p)^{b'-1}$$

$$= \mathrm{Beta}(a', b')$$

$$= \mathrm{Beta}(a+k, b+n-k)$$

$$= \mathrm{Beta}(a', b')$$

$$= \mathrm{Beta}(k+1, n-k+1) \qquad a = 1,\ b=1$$

$$E(p) = \frac{k+1}{n+2}$$

What is the probability of getting a head?

$n = 10, k = 7$

Expected value

- 7/10 = 70%

- How confident are you?
- What if 70 out of 100 reviews were positive?
- What if you knew that it was more likely for the data to tampered?

What is the probability of getting a head?

Before

$n = 10, k = 7$

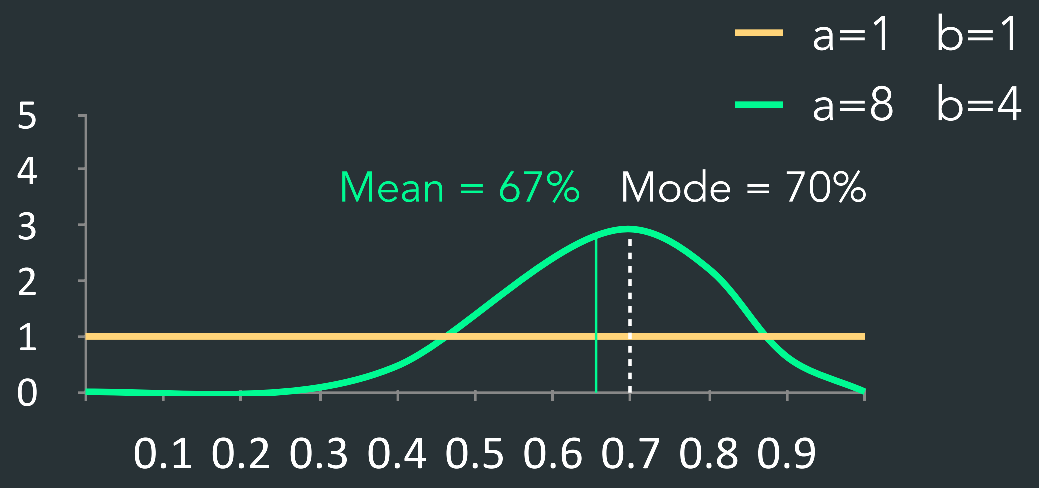
Expected value

- $7/10 = 70\%$

- How confident are you?
- What if 70 out of 100 reviews were positive?
- What if you knew that it was more likely for the data to tampered?

After

- Start with Beta(1, 1)
- Update prior belief with Beta(1+7, 1+10-7)
- Expected value
  $= 8/12 = 67\%$



Takeaways

- Recap: Binomial distribution
- Continuous parameters / hypotheses can be estimated using Bayes Rule
- Conjugate priors allow us to analytically find posterior distributions
- Beta distribution is a conjugate prior for the Binomial distribution
- Estimate the probability of success of a coin flip using Bayes Rule

Bernoulli Distribution

- $Y \sim$ Bernoulli(p)
  - $X = 0, 1$
  - $0 < p < 1$
- Events with 1 Trial & 2 Possible Outcomes
- Examples
  - 1 coin flip
  - Answers to True/False Quiz
  - Voting in a 2-Candidate Election
  - Result of a COVID-19 Test