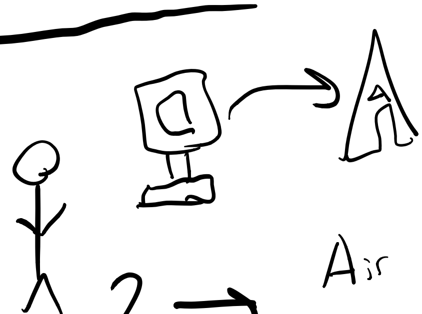
Data 102 Lecture 6: Bayesian Inference

Topics for today:

- Review of Bayesian probability
- Forms of inference
- Start on graphical models

Bayesian modeling.



2 → Air Purifier #1
5 reviews: all positive (5 star)

40 → Air Purifier #2
20 reviews: 19 positive (5 star)
1 negative (2 star)

θ_1, θ_2

θ_i: probability that product i gets positive review

observe: actual reviews

Focus on product 1

negative   positive

Observe $x_1, \dots, x_5 \in \{0, 1\}$

Unknown $\theta \in [0,1]$: prob. pos

Likelihood

$$p(x \mid \theta)$$

$$p(x_1 = 1 \mid \theta) = \theta$$

$$p(x_1 = 0 \mid \theta) = 1 - \theta$$

$$p(x_1 \mid \theta) = \theta^{x_1} (1-\theta)^{1-x_1}$$

x's appear in exponent

✓ $x_1 = 1$: $\theta^1 (1-\theta)^{\cancel{0}}$ $= \theta$

✓ $x_1 = 0$: $\theta^0 (1-\theta)^1 = 1 - \theta$

$$\underbrace{p(x_1, \dots, x_5 \mid \theta)}_{\text{likelihood function}} = \theta^{x_1}(1-\theta)^{1-x_1} \theta^{x_2} (1-\theta)^{1-x_2} \cdots \theta^{x_5} (1-\theta)^{1-x_5}$$

$$= \theta^{x_1 + x_2 + \dots + x_5} (1-\theta)^{5 - (x_1 + x_2 + \dots + x_5)}$$

$$= \theta^{\# \text{pos}} (1-\theta)^{\# \text{negative}}$$

$$x_1 := \begin{cases} 1: & \text{customer \# 1 has positive review} \\ 0: & \text{customer \# 1 has negative review} \end{cases}$$

$$p(x_1 | \theta) = \begin{cases} \theta : & x_1 = 1 \\ 1-\theta : & x_1 = 0 \end{cases}$$

Bernoulli random variable $= \theta^{x_1}(1-\theta)^{1-x_1}$

Goal: Say something about θ

Starting point: MLE (maximum likelihood estimation)

Product #1:

$$\hat{\theta}_{MLE} = \underset{\theta}{\arg\max}\ p(x_1, \dots, x_5 | \theta)$$

$$= \theta^{\# pos} (1-\theta)^{\# neg}$$

$(x_1, \dots, x_5$ all equal $1)$

$$= \underset{\theta}{\arg\max}\ \theta^5 \cancel{(1-\theta)}^0$$

$$= 1$$

Problem #2:

$$\hat{\theta}_{MLE} = \underset{\theta}{\operatorname{argmax}}\ \theta^{19}(1-\theta)^{1}$$

$$= \underset{\theta}{\operatorname{argmax}}\ \log\left(\theta^{19}(1-\theta)^{1}\right)$$

$$= \underset{\theta}{\operatorname{argmax}}\ 19\log(\theta) + \log(1-\theta)$$

$$= 19/20$$

MLE: maximizer (over θ) of $p(x|\theta)$

Bayesian approach:

Care about θ

$$p(\theta|x) = \frac{p(x|\theta)\,p(\theta)}{p(x)}$$

$$p(\theta|x_1,\dots,x_5) = \frac{p(x_1,\dots,x_5|\theta)\cdot p(\theta)}{p(x_1,\dots,x_5)}$$

posterior

know

prior

(likelihood) don't care

$$p(\theta|x_1,\dots,x_5) \propto p(x_1,\dots,x_5|\theta) \cdot p(\theta)$$

"is proportional to"

Two questions:

- What is prior $p(\theta)$?
- How should I make my decisions?

Example prior: Beta Distribution

$$p(\theta) \propto \theta^{r-1}(1-\theta)^{s-1}$$

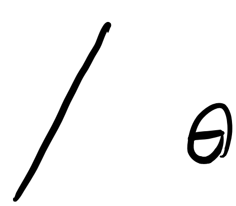
$(r, s > 0)$

Beta(r,s)

Beta$(1,1)$ uniform over $[0,1]$

$$\theta^{1-1}(1-\theta)^{1-1} = 1$$

Beta$(2,1)$



(skew toward 1)

$\mathrm{Beta}(1,2)$ $\quad 1-\theta \quad$ (skew toward 0)

"conjugate prior" of Bernoulli

$$p(\theta) = \theta^{r-1}(1-\theta)^{s-1}$$

$$p(\theta \mid \underbrace{x_1,\dots,x_5}_{\text{all }=1}) \propto \underbrace{p(x_1,\dots,x_5 \mid \theta)}_{\theta^5} \; \underbrace{p(\theta)}_{\theta^{r-1}(1-\theta)^{s-1}}$$

$$= \theta^{r+5-1}(1-\theta)^{s-1}$$

$$= \mathrm{Beta}(r+5, s)$$

Continuous data:

$$\underbrace{x_1,\dots,x_{100}}_{\text{heights}} \in \mathbb{R}_{\geq 0}$$

Θ : average height

$$p(x_1|\theta) \propto \exp\left(-\frac{1}{2\sigma^2}(x_1-\theta)^2\right)$$

Gaussian with mean θ

(assume σ is known)

$$p(\theta) \propto \exp\left(-\frac{1}{2\sigma_2^2}(\theta-\mu)^2\right)$$

Decisions:

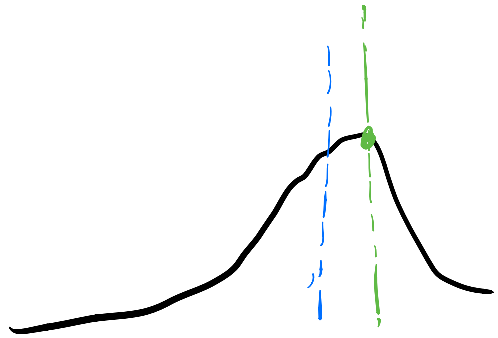
- Need to actually choose a product to buy.
- Ways to do this:
  - Define a loss function
  - Provide some way of

Summarizing $p(\theta|x)$, decide based on that

- posterior mean

  $$\mathbb{E}[\theta|X]$$

  (MLSE)



- posterior mode

  a.k.a. maximum a posteriori

  (MAP)

Computation

Approximate inference

Inference via sampling

Lecture 8