

Unit 2: Probability and distributions

2. Bayes' theorem and Bayesian inference

Sta 101 - Fall 2015

Duke University, Department of Statistical Science

Dr. Çetinkaya-Rundel

Slides posted at http://bit.ly/sta101_f15

1

1. Probability trees are useful for conditional probability calculations

- ▶ Probability trees are useful for organizing information in conditional probability calculations
- ▶ They're especially useful in cases where you know $P(A | B)$, along with some other information, and you're asked for $P(B | A)$

2

- ▶ If you received an email from me about your clicker registration being missing and you still have not given me your info on the Google doc, please do that ASAP!
- ▶ PS2 is posted
- ▶ Lab 1 is due tomorrow before the beginning of your lab section

2. Bayesian inference: start with a prior, collect data, calculate posterior, make a decision or iterate

- ▶ In Bayesian inference, probabilities are at times interpreted as **degrees of belief**.
- ▶ You start with a set of **prior beliefs** (or prior probabilities).
- ▶ You observe some data.
- ▶ Based on that data, you update your beliefs.
- ▶ These new beliefs are called **posterior beliefs** (or posterior probabilities), because they are **post**-data.
- ▶ You can iterate this process.

3

We'll play a game to demonstrate this approach:

- ▶ Two dice: 6-sided and 12-sided
 - I keep one die on the left and one die on the right
- ▶ The “good die” is the 12-sided die.
- ▶ Ultimate goal: come to a class consensus about whether the die on the left or the die on the right is the “good die”
- ▶ We will start with priors, collect data, and calculate posteriors, and make a decision or iterate until we're ready to make a decision

4

- ▶ At each roll I tell you whether you won or not ($\text{win} = \geq 4$)
 - $P(\text{win} \mid \text{6-sided die}) = 0.5 \rightarrow \text{bad die}$
 - $P(\text{win} \mid \text{12-sided die}) = 0.75 \rightarrow \text{good die}$
- ▶ The two competing claims are
 - H_1 : Good die is on left
 - H_2 : Good die is on right
- ▶ Since initially you have no idea which is true, you can assign equal *prior probabilities* to the hypotheses
 - $P(H_1 \text{ is true}) = 0.5$
 - $P(H_2 \text{ is true}) = 0.5$

5

Rules of the game

- ▶ You won't know which die I'm holding in which hand, left (L) or right (R). left = YOUR left
- ▶ You pick die (L or R), I roll it, and I tell you if you win or not, where winning is getting a number ≥ 4 . If you win, you get a piece of candy. If you lose, I get to keep the candy.
- ▶ We'll play this multiple times with different contestants.
- ▶ I will not swap the sides the dice are on at any point.
- ▶ You get to pick how long you want play, but there are costs associated with playing longer.

6

Hypotheses and decisions

Decision	Truth	
	L good, R bad	L bad, R good
Pick L	You get candy!	You lose all the candy :(
Pick R	You lose all the candy :(You get candy!

Sampling isn't free!

At each trial you risk losing pieces of candy if you lose (the die comes up < 4). Too many trials means you won't have much candy left. And if we spend too much class time and we may not get through all the material.

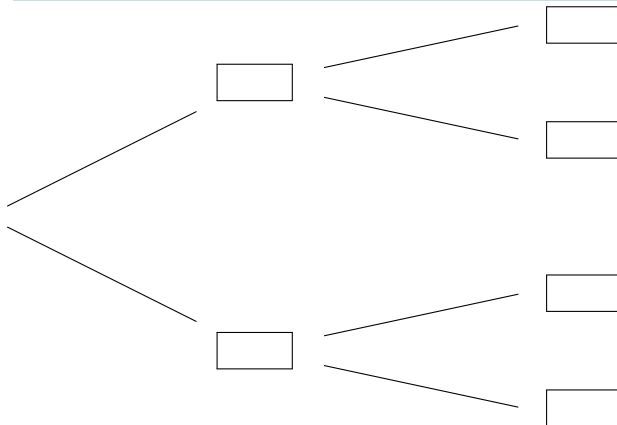
7

	Choice (L or R)	Result (win or loss)
Roll 1		
Roll 2		
Roll 3		
Roll 4		
Roll 5		
Roll 6		
Roll 7		
...		

What is your decision? How did you make this decision?

8

Calculate the posterior probability for the hypothesis chosen in the first roll, and discuss how this might influence your decision for the next roll.



10

- *Posterior probability* is the probability of the hypothesis given the observed data: $P(\text{hypothesis} \mid \text{data})$
- Using Bayes' theorem

$$\begin{aligned}
 P(\text{hypothesis} \mid \text{data}) &= \frac{P(\text{hypothesis and data})}{P(\text{data})} \\
 &= \frac{P(\text{data} \mid \text{hypothesis}) \times P(\text{hypothesis})}{P(\text{data})}
 \end{aligned}$$

9

3. Posterior probability and p-value do not mean the same thing

- *p-value* : $P(\text{observed or more extreme outcome} \mid \text{null hypothesis is true})$
 - This is more like $P(\text{data} \mid \text{hyp})$ than $P(\text{hyp} \mid \text{data})$.
- *posterior* : $P(\text{hypothesis} \mid \text{data})$
- Bayesian approach avoids the counter-intuitive Frequentist p-value for decision making, and more advanced Bayesian techniques offer flexibility not present in Frequentist models
- *Watch out!*
 - *Bayes*: A good prior helps, a bad prior hurts, but the prior matters less the more data you have.
 - *p-value*: It is really easy to mess up p-values: [Goodman, 2008](#)

11

Application exercise: 2.2 Bayesian inference for drug testing

See the [course website](#) for instructions.

1. Probability trees are useful for conditional probability calculations
2. Bayesian inference: start with a prior, collect data, calculate posterior, make a decision or iterate
3. Posterior probability and p-value do not mean the same thing