Unit 2: Probability and distributions

2. Bayes' theorem and Bayesian inference

Sta 101 - Spring 2015

Duke University, Department of Statistical Science

January 28, 2015

Dr. Çetinkaya-Rundel

Slides posted at http://bitly.com/sta101sp15

1. Housekeeping

2. Main ideas
    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

3. Summary

- Review Project 1 assignment and start thinking about data you might want to find / collect for your project

1. Housekeeping

2. Main ideas

    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

3. Summary

1. Housekeeping

2. Main ideas

    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

3. Summary

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)

1. Housekeeping

2. Main ideas

    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

3. Summary

We’ll play a game to demonstrate this approach...

- Two dice: 6-sided and 12-sided
  - I keep one die in my left hand and one die on the right

- Two dice: 6-sided and 12-sided
  - I keep one die in my left hand and one die on the right
- Ultimate goal: come to a class consensus about whether the die on the left or the die on the right is the "good die"

- Two dice: 6-sided and 12-sided
  - I keep one die in my left hand and one die on the right
- 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

- At each roll I tell you whether you won or not (win = ≥ 4)
  - P(win on 6-sided die) = 0.5 $\rightarrow$ bad die
  - P(win on 12-sided die) = 0.75 $\rightarrow$ good die

- At each roll I tell you whether you won or not (win = ≥ 4)
  - P(win on 6-sided die) = 0.5 → bad die
  - P(win on 12-sided die) = 0.75 → good die
- The two competing claims are

  H_1: Good die is in left hand
  H_2: Good die is in right hand

- At each roll I tell you whether you won or not (win = ≥ 4)
  - P(win on 6-sided die) = 0.5 → bad die
  - P(win on 12-sided die) = 0.75 → good die
- The two competing claims are

  H_1: Good die is in left hand
  H_2: Good die is in right hand
- Since initially you have no idea which is true, you can assign equal *prior probabilities* to the hypotheses

  P(H_1 is true) = 0.5
  P(H_2 is true) = 0.5

- You won’t know which die I’m holding in which hand, left (L) or right (R). left = YOUR left

- 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.

- 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.

- 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 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.

	Truth	
Decision	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.

	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?

- *Posterior probability* is the probability of the hypothesis given the observed data: P(hypothesis | data)

- *Posterior probability* is the probability of the hypothesis given the observed data: P(hypothesis | data)
- Using Bayes' theorem

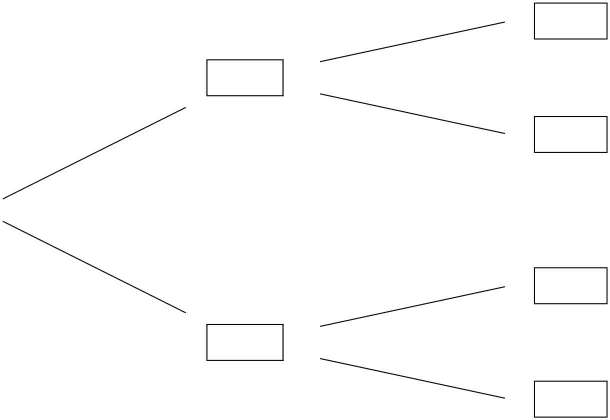
$$\mathrm{P}(\text{hypothesis data}) \quad = \quad \frac{\mathrm{P}(\text{hypothesis and data})}{\mathrm{P}(\text{data})}$$

Posterior probability

- *Posterior probability* is the probability of the hypothesis given the observed data: P(hypothesis | data)
- Using Bayes' theorem

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

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



1. Housekeeping

2. Main ideas

   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

3. Summary

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

- *p-value* : P(observed or more extreme outcome | null hypothesis is true)
  - This is roughly P(data | hypothesis)
- *posterior* : P(hypothesis | 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!*
  A good prior helps, a bad prior hurts, but the prior matters less the more data you have.

Application exercise: 2.2 Bayesian inference for drug testing

See the course website for instructions.

1. Housekeeping

2. Main ideas

    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

3. Summary

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