Unit 2: Probability and distributions

2. Bayes' theorem and Bayesian inference

Sta 104 - Summer 2015

Duke University, Department of Statistical Science

May 20, 2015

Dr. Çetinkaya-Rundel

Slides posted at http://bit.ly/sta104su15

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

        Dice Game

    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
- Think carefully about what the population is and the cases in the population. Do note use summary statistics for your data set.

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

        Dice Game

    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

        Dice Game

    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 $P(A \mid B)$, along with some other information, and you're asked for $P(B \mid A)$

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

        Dice Game

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

3. Summary

- In Bayesian inference, probabilities are at times interpreted as **degrees of belief**.

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).

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.

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.

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 beleifs are called **posterior beliefs** (or posterior probabilities), because they are **post**-data.

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 beleifs are called **posterior beliefs** (or posterior probabilities), because they are **post**-data.
- You can iterate this process.

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

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

- 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.

- 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”

- 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

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

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

  H_1: Good die is on left

  H_2: Good die is on right

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

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.

	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)

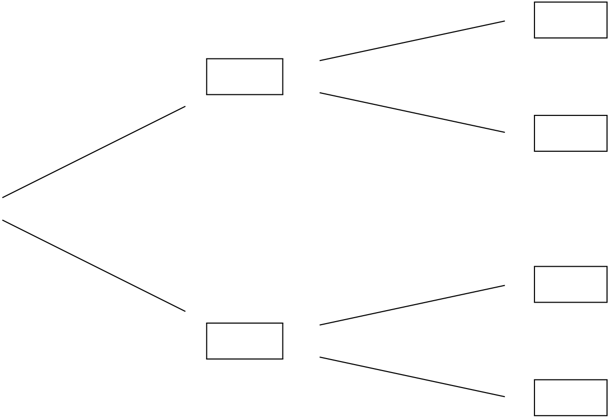
- *Posterior probability* is the probability of the hypothesis given the observed data: P(hypothesis | data)
- Using Bayes' theorem

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

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

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

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
        Dice Game
    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 more like P(data | hyp) than P(hyp | data).
- *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!*
  - *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

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

        Dice Game

    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