# All your Bayes are belong to us: Puzzles in Conditional Probability

Peter Zoogman
Jacob Group Graduate Student Forum
June 5, 2013

# Bayesian Analysis – not very intuitive!

- Humans are bad at combining *a priori* information with new observations

- Very important applications in a variety of fields…and everyday life!
  - Atmospheric data assimilations (duh)
  - Interpreting medical/criminal test results
  - Job interview questions!

  - Understanding (or questioning) safety information:

# What does a positive really mean?

- You are tested for a medical condition that occurs in 1 in 1000 people
- You are told that the test does not produce false negatives, and only produces false positives 5% of the time
- You test positive – what is the probability you have the condition?

This problem was given to students/staff at Harvard Med school in the '70s – how do you think they did?

# What does a positive really mean?

- You are tested for a medical condition that occurs in 1 in 1000 people
- You are told that the test does not produce false negatives, and only produces false positives 5% of the time
- You test positive – what is the probability you have the condition?

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

*a priori* knowledge

$$P(condition|positive) = \frac{P(positive|condition)P(condition)}{P(positive)}$$

# What does a positive really mean?

- You are tested for a medical condition that occurs in 1 in 1000 people
- You are told that the test does not produce false negatives, and only produces false positives 5% of the time
- You test positive – what is the probability you have the condition?

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

*a priori* knowledge

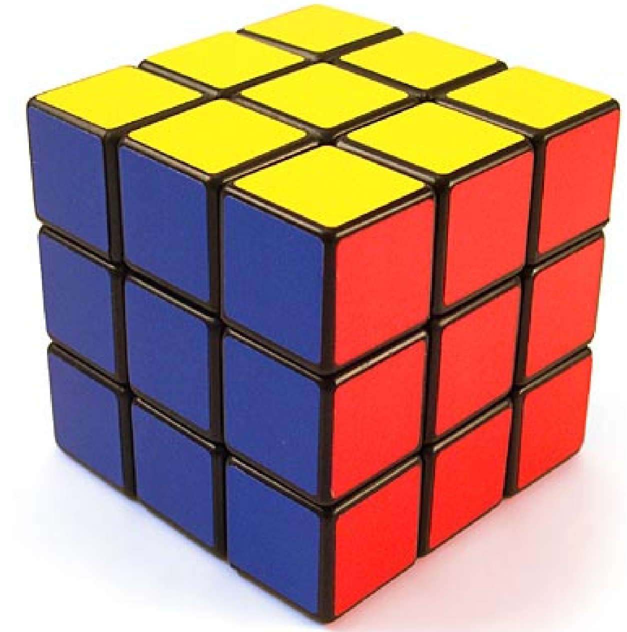
$$P(condition|positive) = \frac{P(positive|condition)P(condition)}{P(positive)}$$

$$P(condition|positive) = \frac{1 * 0.001}{0.001 + .05 * .999} = \sim 2\%$$

- When this problem was given to Harvard Med school students – over 50% responded with 95%, only ~18% with the correct answer [Casscells et al. 1978]

# Adjusting Probabilities w/ Observations

- Start with a cube that is painted on its exterior and then divided into 27 equal cubes
- Put the 27 cubes in a bag, draw one at random and without looking place it on a table in front of you
  - If none of the (5) sides you can see have paint on them, what is the probability the face down side is painted?



# Adjusting Probabilities w/ Observations

- Start with a cube that is painted on its exterior and then divided into 27 equal cubes
- Put the 27 cubes in a bag, draw one at random and without looking place it on a table in front of you
  - If none of the (5) sides you can see have paint on them, what is the probability the face down side is painted?

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Observational information

$$P(painted\ side\ down|no\ paint\ visible) = \frac{P(no\ paint\ vis|painted\ side\ down)P(paint)}{P(no\ paint\ vis)}$$

# Adjusting Probabilities w/ Observations

- Start with a cube that is painted on its exterior and then divided into 27 equal cubes
- Put the 27 cubes in a bag, draw one at random and without looking place it on a table in front of you
  - If none of the (5) sides you can see have paint on them, what is the probability the face down side is painted?

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$P(painted\ side\ down|no\ paint\ visible) = \frac{P(no\ paint\ vis|painted\ side\ down)P(paint)}{P(no\ paint\ vis)}$$

$$P(1\ side\ painted|no\ paint\ visible) = \frac{P(no\ paint\ vis|1\ side\ painted)P(1\ side\ painted)}{P(no\ paint\ vis)}$$

$$P(1\ side\ painted|no\ paint\ visible) = \frac{\frac{1}{6} * \frac{6}{7}}{\frac{2}{7}} = \frac{1}{2}$$

# How far away?

- You are the captain of a ship near shore – two crewmen report that you are distances $d_1$=300 m and $d_2$=100 m away from shore (or ice!).
- You know that they have vision of 20/20 and 20/30 respectively (giving some standard deviations $s_1$ and $s_2$)
- What is the best estimate of your distance from shore?

Bob Pickart, WHOI

# How far away?

- You are the captain of a ship near shore – two crewmen report that you are distances $d_1$=300 m and $d_2$=100 m away from shore (or ice!).
- You know that they have vision of 20/20 and 20/30 respectively (giving some standard deviations $s_1$ and $s_2$)
- What is the best estimate of your distance from shore?

$$d_{opt} = \mathrm{A}d_1 + (1 - \mathrm{A})d_2$$

$$Var(d_{opt}) = \mathrm{A}^2 {s_1}^2 + (1 - \mathrm{A})^2 {s_2}^2$$

# How far away?

- You are the captain of a ship near shore – two crewmen report that you are distances $d_1$=300 m and $d_2$=100 m away from shore (or ice!).
- You know that they have vision of 20/20 and 20/30 respectively (giving some standard deviations $s_1$ and $s_2$)
- What is the best estimate of your distance from shore?

$$d_{opt} = \mathrm{A}d_1 + (1 - \mathrm{A})d_2$$

$$Var(d_{opt}) = \mathrm{A}^2 {s_1}^2 + (1 - \mathrm{A})^2 {s_2}^2$$

$$\frac{\partial Var(d_{opt})}{\partial A} = 2\mathrm{A}{s_1}^2 + 2(1 - \mathrm{A}){s_2}^2 = 0$$

# How far away?

- You are the captain of a ship near shore – two crewmen report that you are distances $d_1$=300 m and $d_2$=100 m away from shore (or ice!).
- You know that they have vision of 20/20 and 20/30 respectively (giving some standard deviations $s_1$ and $s_2$)
- What is the best estimate of your distance from shore?

$$d_{opt} = \mathrm{A}d_1 + (1 - \mathrm{A})d_2$$

$$Var(d_{opt}) = \mathrm{A}^2{s_1}^2 + (1 - \mathrm{A})^2{s_2}^2$$

$$\frac{\partial Var(d_{opt})}{\partial A} = 2\mathrm{A}{s_1}^2 + 2(1 - \mathrm{A}){s_2}^2 = 0$$

$$A = \frac{{s_2}^2}{{s_1}^2 + {s_2}^2} \sim 0.69$$

$$d_{opt} = \mathrm{A}d_1 + (1 - \mathrm{A})d_2 = 228\, m$$

Optimal interpolation!

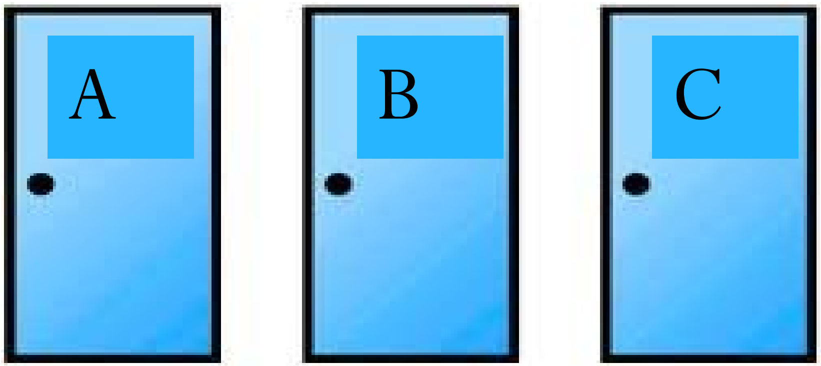
# The Girl Named Florida

- If a family has two children, what is the chance that they have two girls?
- If a family has two children, and at least one of them is a girl, what is the chance that they have two girls?
- If a family has two children, and the oldest one is a girl, what is the chance that they have two girls?
- If a family has two children, and one of them is a girl named Florida, what is the chance that they have two girls?

# The Monty Hall Problem

You're given the choice of three doors:

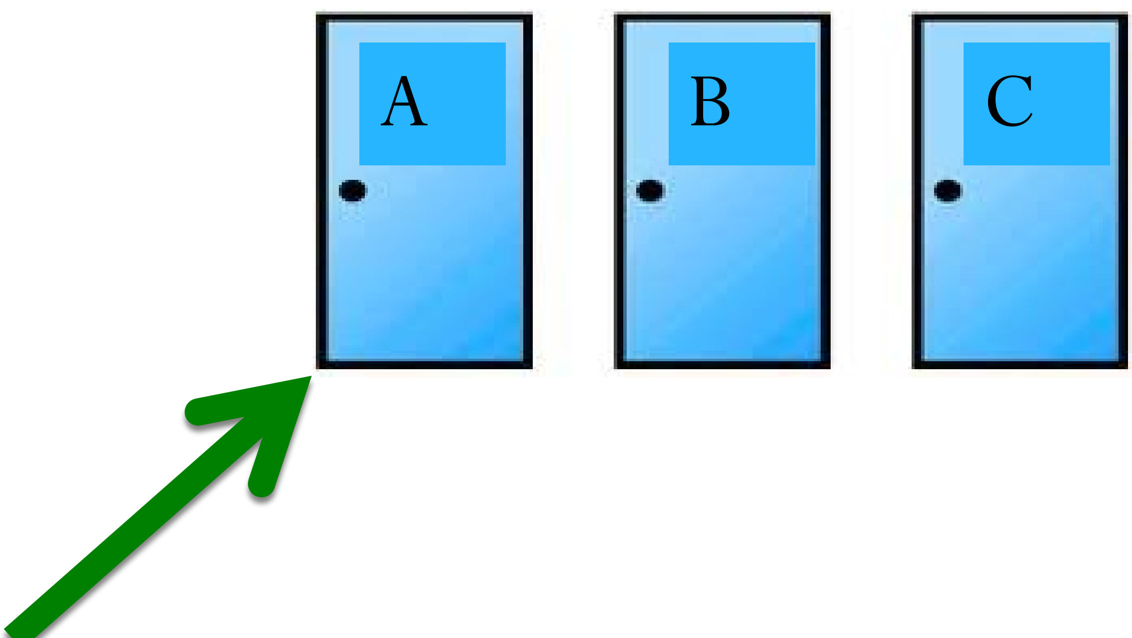
Behind one door is a car; behind the others, goats. You pick a door, say door A [but the door is not opened], and the host, who knows what's behind the doors, opens another door, say door B, which has a goat. He then says to you, "Do you want to pick door C?" Is it to your advantage to switch your choice?



# The Monty Hall Problem

You're given the choice of three doors:

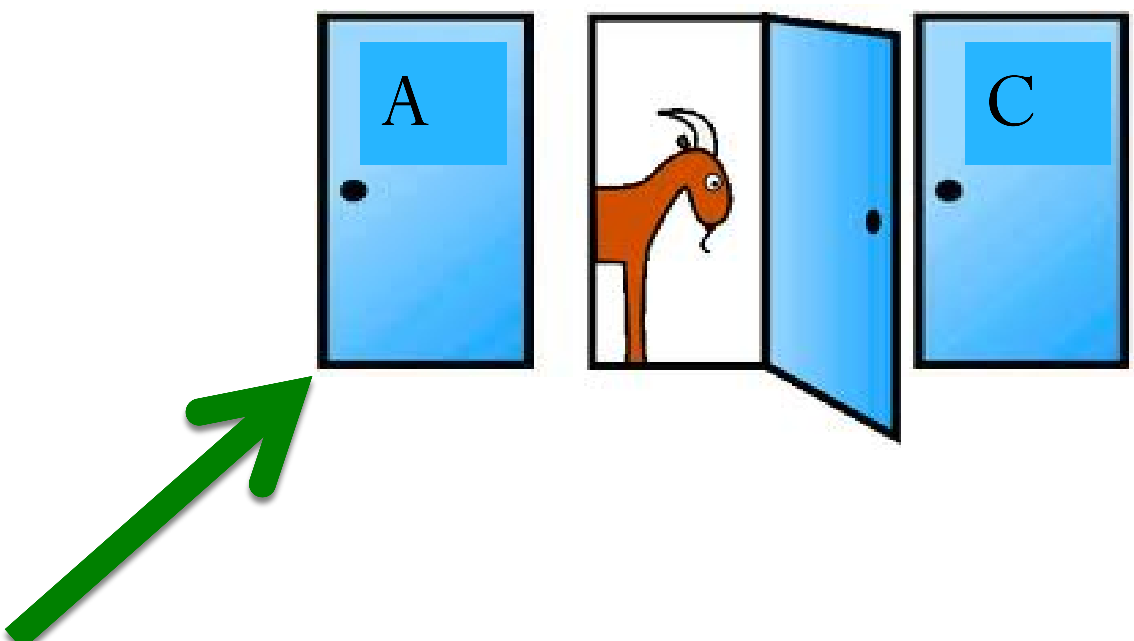
Behind one door is a car; behind the others, goats. You pick a door, say door A [but the door is not opened], and the host, who knows what's behind the doors, opens another door, say door B, which has a goat. He then says to you, "Do you want to pick door C?" Is it to your advantage to switch your choice?



# The Monty Hall Problem

You're given the choice of three doors:

Behind one door is a car; behind the others, goats. You pick a door, say door A [but the door is not opened], and the host, who knows what's behind the doors, opens another door, say door B, which has a goat. He then says to you, "Do you want to pick door C?" Is it to your advantage to switch your choice?



# Bayes' Theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$P(\text{car is behind A}|\text{door B opened}) = \frac{P(\text{door B opened}|\text{car is behind A})P(\text{car is behind A})}{P(\text{door B opened})}$$

$$P(\text{door B opened}|\text{car is behind A}) = 0.5$$

$$P(\text{car is behind A}|\text{door B opened}) = \frac{0.5 * 0.33}{0.5} = 0.33$$

Why is door C different?

# Bayes' Theorem

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

$$P(\text{car is behind A} \mid \text{door B opened}) = \frac{P(\text{door B opened} \mid \text{car is behind A})P(\text{car is behind A})}{P(\text{door B opened})}$$

$$P(\text{door B opened} \mid \text{car is behind A}) = 0.5$$

$$P(\text{car is behind A} \mid \text{door B opened}) = \frac{0.5 * 0.33}{0.5} = 0.33$$

Why is door C different?

$$\boxed{P(\text{door B opened} \mid \text{car is behind C}) = 1}$$

# Updating an Initial Guess

- A cab was involved in a hit-and-run accident at night
  - 85% of the cabs in the city are Green and 15% are Blue
  - A witness identified the cab as Blue. Tests show that this witness correctly identifies the color of the cab 80% of the time

What is the probability the cab was involved in the accident was Blue?



Set-up from Thinking, Fast and Slow (Kahneman)

# Combine observation with a priori information

$$P(\text{car is Blue} \mid \text{witnessed Blue}) = \frac{P(\text{witnessed Blue} \mid \text{car is Blue})P(\text{car is blue})}{P(\text{witnessed Blue})}$$

$$P(\text{car is blue}) = 0.15$$

$$P(\text{witnessed Blue} \mid \text{car is Blue}) = 0.8$$

$$P(\text{witnessed Blue}) = 0.8 * 0.15 + 0.2 * 0.85 = 0.29$$

$$P(\text{car is Blue} \mid \text{witnessed Blue}) = \frac{0.8 * 0.15}{0.29} = 0.41$$