Lecture Stat 302
Introduction to Probability - Slides 5

AD

Jan. 2010

Conditional Probabilities

- **Conditional Probability**. Consider an experiment with sample space S. Let E and F be two events, then the conditional probability of E given F is denoted by $P(E|F)$ and satisfies if $P(F) > 0$
$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$
- **Intuition**: If F has occured, then, in order for E to occur, it is necessary that the occurence be both in E and F, hence it must be in $E \cap F$. Once F has occured, F is the new sample space.
- *Equally likely outcomes*. In this case, we have

$$\begin{aligned} P(E|F) &= \frac{\#\text{ outcomes in } E \cap F}{\#\text{ outcomes in } F} \\ &= \underbrace{\frac{\#\text{ outcomes in } E \cap F}{\#\text{ outcomes in } S}}_{P(E \cap F)} / \underbrace{\left(\frac{\#\text{ outcomes in } F}{\#\text{ outcomes in } S}\right)}_{P(F)}. \end{aligned}$$

Example

- Mr and Mrs. Smith have 2 children, boys or girls, the 4 configurations being equally likely. What is the probability that the two children are two girls if **1)** you are not given any additional information, **2)** the elder child is a girl, **3)** one of the two children is a girl.
- **Answer to 1)**: By assumption, we have $P(\{G,G\}) = P(\{G,B\}) = P(\{B,G\}) = P(\{B,B\}) = \frac{1}{4}$ so if you don't have any additional information $P(\{G,G\}) = 1$.
- **Answer to 2)**: Now we know that the elder child is a girl so the event $F = (\{G,G\},\{G,B\})$ occured. We are interested in

$$P(\{G,G\}|\, F) = \frac{P(F \cap \{G,G\})}{P(F)} = \frac{P(\{G,G\})}{P(F)} = \frac{\frac{1}{4}}{\frac{2}{4}} = \frac{1}{2}.$$

- **Answer to 3):** Now we know that one of the two children is a girl so the event $F = (\{G,G\},\{G,B\},\{B,G\})$ occured. We have

$$P(\{G,G\}|\, F) = \frac{P(F \cap \{G,G\})}{P(F)} = \frac{P(\{G,G\})}{P(F)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}.$$

Examples

- **Example**: A math teacher gave her class two tests. 25% of the class passed both tests and 42% of the class passed the first test. What percent of those who passed the first test also passed the second test?
- **Answer**: Let $E_i =$ "Event passed test i". We have

$$P(E_1) = 0.42 \text{ and } P(E_1 \cap E_2) = 0.25$$

so

$$P(E_2 | E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)} = \frac{0.25}{0.42} = 0.60$$

- **Example**: The probability that it is Friday and that a student is absent is 0.03. Since there are 5 school days in a week, the probability that a school day is Friday is 0.2. What is the probability that a student is absent given that today is Friday?
- **Answer**: We have

$$P(\text{Absent} | \text{Friday}) = \frac{P(\text{Absent on Friday})}{P(\text{Friday})} = \frac{0.03}{0.2} = 0.15.$$

The Multiplication Rule

- Let $E_1, E_2, \ldots, E_n$ be a sequence of events, then we have

$$P(E_1 \cap E_2 \cap \cdots \cap E_n) = P(E_1)\, P(E_2 | E_1) \times$$
$$\times P(E_3 | E_1 \cap E_2) \cdots P(E_n | E_1 \cap \cdots \cap E_{n-1})$$

- Proof follows from

$$P(E_1 \cap \cdots \cap E_n) = \tfrac{P(E_1 \cap \cdots \cap E_n)}{P(E_1 \cap \cdots \cap E_{n-1})} P(E_1 \cap E_2 \cap \cdots \cap E_{n-1})$$
$$= P(E_n | E_1 \cap \cdots \cap E_{n-1}) \tfrac{P(E_1 \cap \cdots \cap E_{n-1})}{P(E_1 \cap \cdots \cap E_{n-2})} P(E_1 \cap \cdots \cap E_{n-2})$$

i.e. apply repeatedly $P(A | B) = P(A \cap B) / P(B)$ where $A = E_1 \cap \cdots \cap E_i$ and $B = E_1 \cap \cdots \cap E_{i-1}$.

Example

- **Example**: You have a box with 3 blue marbles, 2 red marbles, and 4 yellow marbles. You are going to pull out one marble, record its color, put it back in the box and draw another marble. What is the proba of pulling out a red followed by a blue?
- **Answer**: We have $P(\text{red}) = 2/9$ and $P(\text{blue}|\text{red first}) = P(\text{blue}) = 3/9$ so

$$P(\text{red} \cap \text{blue}) = P(\text{red})\, P(\text{blue}|\text{red first}) = \frac{2}{9} \times \frac{3}{9} = \frac{2}{27}.$$

- **Example**: Consider the same box of marbles. However, we are going to pull out the first marble, leave it out and then pull out the second marble. What is the probability of pulling out a red marble followed by a blue marble?
- **Answer**: We have $P(\text{red}) = 2/9$ and $P(\text{blue}|\text{red first}) = 3/8 \neq P(\text{blue})$ so

$$P(\text{red} \cap \text{blue}) = P(\text{red})\, P(\text{blue}|\text{red first}) = \frac{2}{9} \times \frac{3}{8} = \frac{1}{12}.$$

Bayes' formula

- **Bayes' formula**. We have for any events E, F

$$\begin{aligned} P(E) &= P(E \cap F) + P(E \cap F^c) \\ &= P(E|F)P(F) + P(E|F^c)P(F^c) \\ &= P(E|F)P(F) + P(E|F^c)(1 - P(F)) \end{aligned}$$

- This can be generalized to the following case. Assume $\cup_{i=1}^{n} F_i = S$ and $F_i \cap F_j = \varnothing$ then

$$\begin{aligned} P(E) &= P(E \cap (\cup_{i=1}^{n} F_i)) = P(\cup_{i=1}^{n} (E \cap F_i)) \\ &= \sum_{i=1}^{n} P(E \cap F_i) \\ &= \sum_{i=1}^{n} P(E|F_i) P(F_i) \end{aligned}$$

Example: Airport Scanner

- A government agency has developed a scanner which determines whether a person is a terrorist. The scanner is fairly reliable; 99% of all scanned terrorists are identified as terrorists, and 95% of all upstanding citizens are identified as such. An informant tells the agency that exactly one passenger of 100 aboard an aeroplane in which in you are seated is a terrorist. The agency decide to scan each passenger, and the shifty looking man sitting next to you tests positive. What are the chances that this man really is a terrorist?
- **Answer.** $A =$"Your neighbour is a terrorist", $B =$ "Your neigbhour tested positive". We are interested in $P(A|B)$ where

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B|A)\,P(A) + P(B|A^c)\,P(A^c)}$$

where $P(A) = \frac{1}{100} = 1 - P(A^c)$, $P(B|A) = 0.99$,
$P(B|A^c) = 1 - 0.95 = 0.05$ so

$$P(A|B) = \frac{0.99/100}{0.99/100 + 0.05 \times 99/100} = 0.1667$$

Example: Screening for HIV test

- HIV blood tests are very accurate: around 99.8% of people with HIV test positive, and 99.99% of people without the virus test negative. In the UK, the prevalence of HIV in adults with no risk factors is around 1 in 10000. For an adult with no risk factors, what is the probability of having HIV given a positive test result? what is the probability of having HIV given a negative test result?
- **Answer.** Consider the event $A =$"Adult has HIV" and $B =$ "HIV test positive". We are interested in $P(A|B)$ where

$$P(A|B) = \frac{P(B|A)\, P(A)}{P(B|A)\, P(A) + P(B|A^c)\, P(A^c)}.$$

Here we have

$$P(A) = \frac{1}{10000} = 1 - P(A^c),\ P(B|A) = 0.998,\ P(B|A^c) = 0.0001$$

so

$$P(A|B) = \frac{0.998/10000}{0.998/10000 + 0.0001 \times 9999/10000} \approx 0.5$$

Example: Screening for HIV test (cont.)

- The probability of having HIV given a negative test result is $P(A|B^c)$ where

$$P(A|B^c) = \frac{P(B^c|A)\,P(A)}{P(B^c|A)\,P(A) + P(B^c|A^c)\,P(A^c)}$$

where

$$P(B^c|A) = 1 - 0.998 = 0.002, \;\; P(B^c|A^c) = 0.9999$$

so

$$P(A|B^c) = \frac{0.002/10000}{0.002/10000 + 0.9999 \times 9999/10000} = 2.10^{-7};$$

there is $2.10^{-5}\%$ chance that the person is HIV positive whereas he/she tested negative.

Example: Screening for Breast Cancer

- Since 1988, women over 50 in the UK have routinely been offered screening for breast cancer, even if they have no other symptoms, and in 2004/05 1.7 million women were screened. Those with a positive mammogram are recalled for further investigations, at considerable cost in anxiety, resources and pain and discomfort. It has been estimated that in women between 50 and 70 years old, mammography will detect approximately 85% of breast cancers. It has been estimated that around 10% of women with no cancer will still receive a positive result. Finally, we need to know something about the underlying incidence of breast cancer in this population. This varies quite a lot by ethnicity and nationality, but in the UK it has been estimated that around 10 out of 1000 women aged 50 to 70 have breast cancer.
- **Question**: For a woman between 50 and 70 years randomly screened, what is the probability of having breast cancer if she tests positive?

Example: Screening for Breast Cancer (cont).

- **Answer**: Consider the event $A =$ "Woman has breast cancer" and $B =$ "Test positive". We are interested in $P(A|B)$ where

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B|A)\,P(A) + P(B|A^c)\,P(A^c)}.$$

Here we have

$$P(A) = 0.01 = 1 - P(A^c),\; P(B|A) = 0.85,\, P(B|A^c) = 0.10$$

so

$$P(A|B) = \frac{0.85 \times 0.01}{0.85 \times 0.01 + 0.10 \times 0.99} \approx 0.08.$$

- There is "only" 8% of chance the woman having breast cancer if the test is positive.

Example: Let's make a deal TV show

- Three doors A,B,C are closed. Behind one of them, there is a Ferrari, behind the others there is a goat. The player selects one door, say A for sake of simplicity. The TV presenter, who knows where the car is, tells him that the car is not behind door B and offers him the possibility to revise her/his choice (i.e. that is to select door C). Should the player revise his choice?
- **Answer**: Let A (resp. B, C) the event "the Ferrari is behind door A" (resp. B, C). Let E the event "the presenter tells the player that it is not behind door B". We want to compute $P(A|E)$. We have

$$P(A) = P(B) = P(C) = \frac{1}{3}$$

and

$$P(E|A) = \frac{1}{2},\ P(E|B) = 0,\ P(E|C) = 1.$$

Indeed, if the car is behind A then the presenter has the choice between B and C. If the car is behind C, the presenter can only pick door B as he/she cannot select A.

Example: Let's make a deal TV show (cont.)

- Hence we have

$$\begin{aligned} P(E) &= P(E|A)\,P(A) + P(E|B)\,P(B) + P(E|C)\,P(C) \\ &= \frac{1}{2} \times \frac{1}{3} + \frac{1}{3} = \frac{1}{2} \end{aligned}$$

  and

$$P(A|E) = \frac{P(E|A)\,P(A)}{P(E)} = \frac{\frac{1}{2} \times \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$$

  and thus

$$P(C|E) = 1 - P(A|E) = \frac{2}{3}.$$

- You should definitely revise your choice.

Example: Prosecutor's Fallacy

- A zealous prosecutor has collected an evidence for a crime and has an expert testify that the proba of finding this evidence if the accused were innocent is one-in-a-million and one if he is guilty. The prosecutor concludes that the probability of the accused being innocent is one-in-a-million. This is WRONG. Compute the probability that the accused is guilty assuming the population of the region where the crime has been committed is of 10 million people
- **Answer**: Define $A =$"The accused is guilty" and $B =$"Finding this evidence" then

$$P(A) = 10^{-7},\ P(B|A) = 1 \text{ and } P(B|A^c) = 10^{-6}$$

so

$$\begin{aligned} P(A|B) &= \frac{P(B|A)\,P(A)}{P(B|A)\,P(A) + P(B|A^c)\,P(A^c)} \\ &= \frac{10^{-7}}{10^{-7} + 10^{-6} \times (1 - 10^{-7})} \approx 0.1. \end{aligned}$$

Example: Sally Clark

- On the 9th November 1999, Sally Clark, then a 35-year-old solicitor was convicted of the murder of her first two children. Her first child, had died three years earlier, aged just under three months. His death had originally been treated as arising by natural causes, probably "Sudden Infant Death Syndrome" (or "SIDS"). Her second child, Harry, died just over a year later at the age of two months. His death was treated as suspicious by the Home Office pathologist who carried out the post mortem. He then revisited Christopher's death and determined that that too was suspicious. Sally Clark was arrested some weeks later and eventually tried at Chester Crown Court.
- The main argument used against her in her trial was that the risk of one SIDS death in a family is 1 in 8,543, hence the risk of two SIDS death is approximately 1 in 73 millions.
- Does it make any sense?

Independence

- We say that two events E, F are independent if

$$P(E \cap F) = P(E)\,P(F)$$

or equivalently

$$P(E|F) = P(E) \text{ and } P(F|E) = P(F).$$

- If E and F are independent then E and F^c are also independent as

$$\begin{aligned} P(E) &= P(E \cap F) + P(E \cap F^c) \\ &= P(E)\,P(F) + P(E \cap F^c) \end{aligned}$$

so

$$P(E \cap F^c) = P(E)(1 - P(F)) = P(E)\,P(F^c).$$

Example

- You toss two air dice. Let E the event that the sum of the dice is 8 and F the event that the first die equals 4. Are E and F independent?
- **Answer.** We have $E = (\{2,6\},\{3,5\},\{4,4\},\{5,3\},\{6,2\})$ and $F = (\{4,1\},\{4,2\},\ldots,\{4,6\})$

$$P(E)\,P(F) = \frac{5}{36} \times \frac{1}{6} = \frac{5}{126}$$

but $E \cap F = \{4,4\}$ and

$$P(E \cap F) = \frac{1}{36}.$$

So the events are not independent.

Example: Game Strategies

- Two players have to cooperate to answer some questions. Both players are equally good. Given a question, they will answer correctly independently with probability p. Two strategies are proposed by the players.
- In Strategy 1, given a question, one selects randomly one of the two players by tossing a fair coin and his/her answer is chosen.
- In Strategy 2, the two players give independently their answers. If their answers agree, then this is the chosen answer. If they disagree, they toss a fair coin and pick randomly one of the answers.
- What is the best strategy?

Example: Game Strategies (cont.)

- Consider the event $R = \{\text{answer is correct}\}$, $P_1 = \{\text{Player 1 answers}\}$, $P_2 = P_1^c = \{\text{Player 2 answers}\}$. In the 1st strategy we have

$$\begin{aligned} P(R) &= P(R|P_1)\,P(P_1) + P(R|P_2)\,P(P_2) \\ &= p \times \frac{1}{2} + p \times \frac{1}{2} = p. \end{aligned}$$

- In the 2nd strategy, we give an correct answer if $A_1 = \{\text{players agree and give correct answer}\}$, $A_2 = \{\text{player 1 right/player 2 wrong and toss decides player 1}\}$ or $A_3 = \{\text{player 1 wrong/player 2 right and toss decides player 2}\}$ occurs. Let A_4 denotes the event where the given answer is incorrect.

$$\begin{aligned} P(R) &= P(R \cap A_1) + P(R \cap A_2) + P(R \cap A_3) + P(R \cap A_4) \\ &= p \times p + p \times (1-p) \times \frac{1}{2} + (1-p) \times p \times \frac{1}{2} = p. \end{aligned}$$

- Strategy 1 and Strategy 2 will give similar results.