
Problem Sheet 3 [1]

Exercise 1. In deterministic logic, the statement "A implies B" is equivalent to the *contrapositive*, "not B implies not A". In this problem we will consider analogous statements in probability, the logic of uncertainty. Let A and B be events with probabilities not equal to 0 or 1.

1. Show that if $P(B|A) = 1$, then $P(A^c|B^c) = 1$

2. Show however that the result in 1. does not hold in general, if "=" is replaced by "≈". In particular, find an example where $P(B|A)$ is very close to 1, but $P(A^c|B^c)$ is very close to 0.

   Hint: What happens if A and B are independent?

Solution 1. 1. Note that $P(A^c|B^c) = 1 - P(A|B^c)$. By Bayes' Rule and the law of total probability

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

If $P(B|A) = 1$, then $P(B^c|A) = 1 - P(B|A) = 0$, thus

$$P(A|B^c) = \frac{0 \cdot P(A)}{0 \cdot P(A) + P(B^c|A^c)P(A^c)} = 0,$$

therefore $P(A^c|B^c) = 1 - 0 = 1$.

2. Suppose that A and B are independent. Then $P(A^c|B^c) = P(A^c)$ and $P(B|A) = P(B) \approx 1$. Further assume that $P(A) \approx 1$, as no assumption on the value of $P(A)$ was made. Then $P(A^c|B^c) = P(A^c) \approx 0$, thus the claim does not hold.

   If we were to assume exact equality and independence i.e. $P(B|A) = P(B) = 1$, then $P(B^c) = 0$, thus $P(A^c|B^c)$ would be left undefined.

Exercise 2. A family has 3 children, creatively named A, B, and C, none of them are twins, i.e. there is a clear ordering in their ages.

1. Discuss intuitively (but clearly) whether the event "A is older than B" is independent of the event "A is older than C".

2. Find the probability that A is older than B, given that A is older than C.

Solution 2. 1. They are not independent: knowing that A is older than B makes it more likely that A is older than C, as if A is older than B, then the only way A can be younger than C is if the birth order is CAB, whereas the birth orders ABC and ACB are both compatible with A being older than B. To make this more intuitive, think of an extreme case where there are 100 children instead of 3, call them $A_1, \ldots, A_{100}$. Given that A_1 is older than all of $A_2, A_3, \ldots, A_{99}$, it's clear that A_1 is very old (relatively), whereas there is no evidence about where A_{100} fits into the birth order.

[1] Exercises are based on the coursebook Statistics 110: Probability by Joe Blitzstein

2. Writing $x > y$ to mean that x is older than y,

$$P(A > B|A > C) = \frac{P(A > B, A > C)}{P(A > C)} = \frac{1/3}{1/2} = \frac{2}{3},$$

since $P(A > B, A > C) = P(A \text{ is the eldest child}) = 1/3$ (unconditionally, any of the 3 children is equally likely to be the eldest), and $P(A > C) = 1/2$, as it is equally likely that A is older than C as it is that C is older than A.

Exercise 3. A family has two children. Let C be a characteristic that a child can have, and assume that each child has characteristic C with probability p, independently of each other and of gender. Under the assumptions of binary gender that have an equal probability for each child ($P(boy) = P(girl)$) and independence between the genders of the children, show that the probability that both children are girls given that at least one is a girl with characteristic C is $\frac{2-p}{4-p}$. Note that this is 1/3 if $p = 1$ and approaches 1/2 from below as $p \to 0$.

Solution 3. By definition

$$P(\text{ both girls}|\text{at least one “characteristic girl”}) = \frac{P(\text{ both girls, at least one “characteristic girl”})}{P(\text{at least one “characteristic girl”})}.$$

The probability that a child is a girl with the characteristic is $1/2 \cdot p$, therefore, the probability that none of the children are girls with the characteristic is $(1 - 1/2 \cdot p)^2$

$$P(\text{at least one “characteristic girl”}) = 1 - (1 - 1/2 \cdot p)^2.$$

Note that the characteristic and the gender are independent, hence

$$\begin{aligned} P(\text{ both girls, at least one “characteristic girl”}) &= P(\text{ both girls, at least one “characteristic child”}) \\ &= P(\text{ both girls})P(\text{at least one “characteristic child”}) \\ &= 1/4 \cdot (1 - (1 - p)^2). \end{aligned}$$

Therefore

$$\begin{aligned} P(\text{ both girls}|\text{at least one “characteristic girl”}) &= \frac{1/4 \cdot (1 - (1 - p)^2)}{1 - (1 - 1/2 \cdot p)^2} \\ &= \frac{1 - 1 + 2p - p^2}{4 - 4 + 4p - p^2} \\ &= \frac{p(2 - p)}{p(4 - p)} \\ &= \frac{2 - p}{4 - p} \end{aligned}$$

Exercise 4. Suppose that there are two types of drivers: good drivers and bad drivers. Let G be the event that a certain person is a good driver, A be the event that they get into a car accident next year, and B be the event that they get into a car accident the following year. Let $P(G) = g$ and $P(A|G) = P(B|G) = p_1$, $P(A|G^c) = P(B|G^c) = p_2$, with $p_1 < p_2$. Suppose that given the information of whether or not the person is a good driver, A and B are independent (for simplicity and to avoid being morbid, assume that the accidents being considered are minor and would not make the person unable to drive).

1. Explain intuitively whether or not A and B are independent.
2. Find $P(G|A^c)$.
3. Find $P(B|A^c)$.

Solution 4. 1. The informal definition of independence of events says, that two events A and B are independent if learning A occurred gives us no information that would change our probabilities for B occurring. Learning that a person had an accident this year (A), reveals information about their driving capabilities (G^c), which consequently changes my perception about the probability of them being in an accident in the following year (B).

   More formally, A and B are independent if $P(AB) = P(A)P(B)$. By the law of total probability and the conditional independence between A and B

$$P(AB) = P(AB|G)P(G) + P(AB|G^c)P(G^c) = \\ P(A|G)P(B|G)P(G) + P(A|G^c)P(B|G^c)P(G^c) = p_1^2 \cdot g + p_2^2 \cdot (1-g).$$

   While by the law of total probability

$$P(A) = P(A|G)P(G) + P(A|G^c)P(G^c) = p_1 \cdot g + p_2 \cdot (1-g),$$

   and as the conditional probabilities are the same for B, $P(A) = P(B)$.

   Therefore

$$P(AB) = p_1^2 \cdot g + p_2^2 \cdot (1-g) \neq (p_1 \cdot g + p_2 \cdot (1-g))^2 = P(A)P(B),$$

   so A and B are not unconditionally independent.

2. By Bayes' rule and the law of total probability

$$P(G|A^c) = \frac{P(A^c|G)P(G)}{P(A^c|G)P(G) + P(A^c|G^c)P(G^c)} = \frac{(1-p_1) \cdot g}{(1-p_1) \cdot g + (1-p_2) \cdot (1-g)}.$$

3. By the law of total probability $P(B|A^c) = P(B|A^c, G)P(G|A^c) + P(B|A^c, G^c)P(G^c|A^c)$. As A and B are conditionally independent given G we have $P(B|A^c, G) = P(B|G)$ and $P(B|A^c, G^c) = P(B|G^c)^2$. Therefore

$$P(B|A^c) = P(B|G)P(G|A^c) + P(B|G^c)P(G^c|A^c) = \\ p_1 \frac{(1-p_1) \cdot g}{(1-p_1) \cdot g + (1-p_2) \cdot (1-g)} + p_2 \frac{(1-p_2) \cdot (1-g)}{(1-p_1) \cdot g + (1-p_2) \cdot (1-g)},$$

   where we used the probability $P(G|A^c)$ derived in part 2.

Exercise 5. You are going to play 2 games of chess with an opponent whom you have never played against before (for the sake of this problem). Your opponent is equally likely to be a beginner, intermediate, or a master. Depending on which, your chances of winning an individual game are 90%, 50%, or 30%, respectively.

[2] $P(B|A^c, G) = \frac{P(B, A^c|G)P(G)}{P(A^c, G)}$ and by conditional independence this is equal to $\frac{P(B|G)P(A^c|G)P(G)}{P(A^c,G)} = \frac{P(B|G)P(A^c,G)}{P(A^c,G)} = P(B|G)$. The same applies for $P(B|A^c, G^c)$.

1. What is your probability of winning the first game?

2. Congratulations: you won the first game! Given this information, what is the probability that you will also win the second game against the same opponent (assume that, given the skill level of your opponent, the outcomes of the games are independent)?

3. Explain the distinction between assuming that the outcomes of the games are independent and assuming that they are conditionally independent given the opponent's skill level. Which of these assumptions seems more reasonable, and why?

Solution 5. 1. Let W_i be the event of winning the i-th game. By the law of total probability,

$$P(W_1) = (0.9 + 0.5 + 0.3)/3 = 17/30.$$

2. We have $P(W_2|W_1) = P(W_2, W_1)/P(W_1)$. The denominator is known from part 1., while the numerator can be found by conditioning on the skill level of the opponent:

$$P(W_1, W_2) = \frac{1}{3}P(W_1, W_2|beginner) + \frac{1}{3}P(W_1, W_2|intermediate) + \frac{1}{3}P(W_1, W_2|expert).$$

Since W_1 and W_2 are conditionally independent given the skill level of the opponent, this becomes

$$P(W_1, W_2) = (0.9^2 + 0.5^2 + 0.3^2)/3 = 23/60.$$

So

$$P(W_2|W_1) = \frac{23/60}{17/30} = 23/34.$$

3. Independence here means that knowing one game's outcome gives no information about the other game's outcome, while conditional independence is the same statement where all probabilities are conditional on the opponent's skill level. Conditional independence given the opponent's skill level is a more reasonable assumption here. This is because winning the first game gives information about the opponent's skill level, which in turn gives information about the result of the second game. That is, if the opponent's skill level is treated as fixed and known, then it may be reasonable to assume independence of games given this information; with the opponent's skill level random, earlier games can be used to help infer the opponent's skill level, which affects the probabilities for future games.

Exercise 6. 1. Suppose that in the population of college applicants, being good at baseball is independent of having a good math score on a certain standardized test (with respect to some measure of "good"). A certain college has a simple admissions procedure: admit an applicant if and only if the applicant is good at baseball or has a good math score on the test.

Give an intuitive explanation of why it makes sense that among students that the college admits, having a good math score is *negatively associated* with being good at baseball, i.e., conditioning on having a good math score decreases the chance of being good at baseball.

2. Show that if A and B are independent and $C = A \cup B$, then A and B are conditionally dependent given C (as long as $P(A \cap B) > 0$ and $P(A \cup B) < 1$), with

$$P(A|B, C) < P(A|C).$$

This phenomenon is known as *Berkson's paradox*, especially in the context of admissions to a school, hospital, etc.

Solution 6. 1. The conditioning set includes both a good mathematics score, and the fact that the student was admitted to a college, and we know that for a successful college application, having either a good math score or being good at baseball is sufficient. If I know someone with a low math score was admitted, then they must be a great baseball player, otherwise, their application would have been rejected. However, the students who were admitted with good math test scores could either be good baseball players or bad baseball players, so knowing that they excel at mathematics, definitely decreases the chance that they are good at baseball.

2. As $C = A \cup B$, it follows that

$$P(A|C) = P(A \cap C)/P(C) = P(A \cap (A \cup B))/P(C) = P(A)/P(C).$$

Similarly,

$$\begin{aligned} P(A|B,C) &= \frac{P(A \cap B \cap C)}{P(B \cap C)} \\ &= \frac{P(A \cap B \cap (A \cup B))}{P(B \cap (A \cup B))} \\ &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{P(B)P(A)}{P(B)} \\ &= P(A), \end{aligned}$$

where in the second to last line we used the fact that A and B are independent. Hence unless $P(C) = P(A \cup B) = 1$, $P(A) < P(A)/P(C)$, however, it is assumed that $P(A \cup B) < 1$, therefore $P(A|B,C) < P(A|C)$. If A and B were conditionally independent given C, then $P(A|B,C)$ would be equal to $P(A|C)$, therefore, we showed that they are not conditionally independent.

Exercise 7. Consider the following conversation from an episode of *The Simpsons*:

> **Lisa**: *Dad, I think he's an ivory dealer! His boots are ivory, his hat is ivory, and I'm pretty sure that check is ivory.*
> **Homer**: *Lisa, a guy who has lots of ivory is less likely to hurt Stampy than a guy whose ivory supplies are low.*

Here Homer and Lisa are debating the question of whether or not the man (named Blackheart) is likely to hurt Stampy the Elephant if they sell Stampy to him. They clearly disagree about how to use their observations about Blackheart to learn about the probability (conditional on the evidence) that Blackheart will hurt Stampy.

1. Define clear notation for the various events (3 in total) of interest here.

2. Express Lisa's and Homer's arguments (Lisa's is partly implicit) as conditional probability statements in terms of your notation from 1.

3. Assume it is true that someone who has a lot of a commodity will have less desire to acquire more of the commodity. Explain what is wrong with Homer's reasoning that the evidence about Blackheart makes it less likely that he will harm Stampy.

Solution 7. 1. Let H be the event that the man will hurt Stampy, let L be the event that a man has lots of ivory, and let D be the event that the man is an ivory dealer.

2. Lisa observes that L is true. She suggests (reasonably) that this evidence makes D more likely, i.e., $P(D \mid L) > P(D)$. Implicitly, she suggests that this makes it likely that the man will hurt Stampy, i.e.,

$$P(H \mid L) > P(H \mid L^c).$$

Homer argues that

$$P(H \mid L) < P(H \mid L^c).$$

3. Homer does not realize that observing that Blackheart has so much ivory makes it much more likely that Blackheart is an ivory dealer, which in turn makes it more likely that the man will hurt Stampy. This is an example of Simpson's paradox. It may be true that, *controlling for whether or not Blackheart is a dealer*, having high ivory supplies makes it less likely that he will harm Stampy: $P(H \mid L, D) < P(H \mid L^c, D)$ and $P(H \mid L, D^c) < P(H \mid L^c, D^c)$. However, this does not imply that $P(H \mid L) < P(H \mid L^c)$.

Exercise 8. Monty Hall is trying out a new version of his game. In this version, instead of there always being 1 car and 2 goats, the prizes behind the doors are generated independently, with each door having probability p of having a car and $q = 1 - p$ of having a goat. In detail: There are three doors, behind each of which there is one prize: either a car or a goat. For each door, there is probability p that there is a car behind it and $q = 1 - p$ that there is a goat, independent of the other doors.

The contestant chooses a door. Monty, who knows the contents of each door, then opens one of the two remaining doors. In choosing which door to open, Monty will always reveal a goat if possible. If both of the remaining doors have the same kind of prize, Monty chooses randomly (with equal probabilities). After opening a door, Monty offers the contestant the option of switching to the other unopened door.

The contestant decides in advance to use the following strategy: first choose door 1. Then, after Monty opens a door, switch to the other unopened door.

1. Find the unconditional probability that the contestant will get a car.

2. **Optional, challenging exercise:** Monty now opens door 2, revealing a goat. Given this information, find the conditional probability that the contestant will get a car.

Solution 8. 1. Let C_i be the event that a car is behind door i, and G_i be the event that a goat is behind door i (technically $C_i = G_i^c$, but I find this notation more readable). A crucial difference compared to the classical Monty Hall problem is that the event $C_i \cap C_j$ for $i \neq j$ has a non-zero probability p^2. From the viewpoint of winning, we are indifferent to what is behind door 1, Monty will never open our chosen door, and we will never leave it to be our final choice, hence we only need to consider doors 2 and 3 (for the sake of completeness, you can check that by conditioning what is behind door 1 the result stays the same). By the law of total probability,

$$\begin{aligned} P(winning) = {} & P(winning|C_2 \cap C_3)P(C_2 \cap C_3) + P(winning|C_2 \cap G_3)P(C_2 \cap G_3) \\ & + P(winning|G_2 \cap C_3)P(G_2 \cap C_3) + P(winning|G_2 \cap G_3)P(G_2 \cap G_3). \end{aligned}$$

Since we employ the switching strategy, we win whenever there is at least one car behind doors 2 and 3, as then there will be a car behind the unopened door, since Monty always opens a goat door if he can. However, if both doors have a goat behind them, we cannot win. Thus, using the independence of the prize generation,

$$P(winning) = 1 \cdot p^2 + 1 \cdot pq + 1 \cdot qp + 0 \cdot q^2 = p^2 + 2qp = p(p + 2q) = p(2 - p).$$

2. Let M_2 be the event that Monty opens door 2 and reveals a goat. Then the probability of interest is $P(C_3|M_2)$, as by the switching strategy we only win if there is a car behind door 3. By the law of total probability

$$\begin{aligned} P(C_3|M_2) = &P(C_3|M_2, C_3, C_2)P(C_3, C_2|M_2) + P(C_3|M_2, G_3, C_2)P(G_3, C_2|M_2) \\ &+ P(C_3|M_2, C_3, G_2)P(C_3, G_2|M_2) + P(C_3|M_2, G_3, G_2)P(G_3, G_2|M_2) \end{aligned}$$

Given that Monty opened door 2 and revealed a goat, the probability that door 2 has a car behind it is 0, hence $P(C_3, C_2|M_2) = P(G_3, C_2|M_2) = 0$. In addition, the probability of having a car behind door 3 given that there is a goat behind it is zero, hence $P(C_3|M_2, G_3, C_2) = P(C_3|M_2, G_3, G_2) = 0$. Thus, the only non-zero term left in the sum is $P(C_3|M_2, C_3, G_2)P(C_3, G_2|M_2)$. The probability of having a car behind door 3, given that there is a car behind it, is straightforwardly 1. For the second term, we use Bayes' rule

$$P(C_3, G_2|M_2) = \frac{P(M_2|C_3, G_2)P(C_3, G_2)}{P(M_2)}.$$

By the law of total probability and independence

$$\begin{aligned} P(M_2) &= P(M_2|C_3, C_2)P(C_3, C_2) + P(M_2|C_3, G_2)P(C_3, G_2) \\ &\quad + P(M_2|G_3, C_2)P(G_3, C_2) + P(M_2|G_3, G_2)P(G_3, G_2) \\ &= 0 \cdot p^2 + 1 \cdot pq + 0 \cdot qp + \frac{1}{2}q^2 = p(1 - p) + \frac{1}{2}(1 - 2p + p^2) = \frac{1}{2}(1 - p)(1 + p). \end{aligned}$$

Therefore,

$$\begin{aligned} P(C_3|M_2) = 1 \cdot P(C_3, G_2|M_2) &= \frac{P(M_2|C_3, G_2)P(C_3, G_2)}{P(M_2)} \\ &= \frac{1 \cdot pq}{1/2 \cdot (1 - p)(1 + p)} = \frac{p(1 - p)}{1/2 \cdot (1 - p)(1 + p)} = \frac{2p}{1 + p}. \end{aligned}$$