

# Introduction to Probability, Fall 2013

## Math 30530 Section 01

### Homework 2 — solutions

1. Chapter 1, problem 14 (“doubles” means that both dice show the same number).

**Solution:** a)  $6/36 = 1/6$ .

b) Sample points in which sum is 4 or less:  $(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (3, 1)$ . Two of these six are doubles, so the required probability is  $2/6 = 1/3$ .

c) Of 36 sample points, 11 have at least one six, so required probability is  $11/36$ .

d) There are 30 sample points in which the two dice land on different numbers. Of these, 10 have at least one six, so required probability is  $10/30 = 1/3$

2. Chapter 1, problem 16.

**Solution:** There are many valid approaches. Here’s one: call the coins  $C_1$  (with heads on both sides; call the sides  $S_1(C_1)$  and  $S_2(C_1)$ ),  $C_2$  (with tails on both sides; call the sides  $S_1(C_2)$  and  $S_2(C_2)$ ), and  $C_3$  (with heads on one side; call this side  $S_1(C_3)$ , and tails on the other side; call this side  $S_2(C_3)$ ).

Our experiment is to choose a coin at random, observe a random side, then observe the other side, and record the pair (what we saw on the first side, what we saw on the second). There are six equally likely possibilities:

Pick  $C_1$ , first observe  $S_1$  — record  $(H, H)$

Pick  $C_1$ , first observe  $S_2$  — record  $(H, H)$

Pick  $C_2$ , first observe  $S_1$  — record  $(T, T)$

Pick  $C_2$ , first observe  $S_2$  — record  $(T, T)$

Pick  $C_3$ , first observe  $S_1$  — record  $(H, T)$

Pick  $C_3$ , first observe  $S_2$  — record  $(T, H)$ .

The *unconditioned* probability that the second face we observe is a tail — i.e., that the opposite face of the randomly chosen coin is tails — is  $3/6 = 1/2$ . *Conditioned* on the first face we observe being a head, we are putting ourselves into a sample space that has three equally points (first, second and fifth above). Of these, only one (the fifth) has tails as the second face observed, so the required conditional probability is  $1/3$ .

Why not  $1/2$ ? (Seeing heads tells us we either picked  $C_1$  or  $C_3$ , each equally likely, one has tails on other side, other doesn’t). The issue is that once we pick  $C_3$  (as likely as

$C_1$ ), we then have to have the toss go right for us to be in the situation that we observe a head on the toss; and this toss goes right only 50% of the time. But for picking  $C_1$ , there's no worries; the toss always goes right. So half the time that we pick  $C_3$ , by tossing it we put ourselves out of the event that we're conditioning on, but we never do this when picking  $C_1$ . Hence, given the conditional information, we're only *half* as likely to have  $C_3$  as  $C_1$ , and the probability of seeing tails on the opposite side is just  $1/3$ .

This is a form of *Bertrand's box paradox*: see [http://en.wikipedia.org/wiki/Bertrand's\\_box\\_paradox#Related\\_problems](http://en.wikipedia.org/wiki/Bertrand's_box_paradox#Related_problems) for much more!

3. Chapter 1, problem 19.

**Solution:** First we look at  $j \neq i$ : the probability that the paper is in drawer  $j$ , given that a search of drawer  $i$  was unsuccessful, is the ratio: (probability that it is in drawer  $j$  AND search of drawer  $i$  unsuccessful) to (probability that search of drawer  $i$  unsuccessful).

We look at the denominator first. The probability that search of drawer  $i$  is successful is  $p_i d_i$  (two things have to happen: it's actually in drawer  $i$ , and her search of drawer  $i$  is successful). So probability that search of drawer  $i$  unsuccessful is  $1 - p_i d_i$ .

Now we look at the numerator. Using the multiplication rule, the probability that it is in drawer  $j$  AND search of drawer  $i$  unsuccessful is the product of  $p_j$  (probability that it is in drawer  $j$ ) and  $1$  (probability search of drawer  $i$  unsuccessful, given that it is in drawer  $j$ ) —  $p_j$  in total.

This gives the claimed probability,  $p_j/(1 - p_i d_i)$ , for the probability that the paper is in drawer  $j$ , given that a search of drawer  $i$  was unsuccessful, when  $j \neq i$ .

Now we look at  $j = i$ : the probability that the paper is in drawer  $i$ , given that a search of drawer  $i$  was unsuccessful, is the ratio: (probability that it is in drawer  $i$  AND search of drawer  $i$  unsuccessful) to (probability that search of drawer  $i$  unsuccessful).

As before the denominator is  $1 - p_i d_i$ .

Now we look at the numerator. Using the multiplication rule, the probability that it is in drawer  $i$  AND search of drawer  $i$  unsuccessful is the product of  $p_i$  (probability that it is in drawer  $i$ ) and  $1 - d_i$  (probability search of drawer  $i$  unsuccessful, given that it is in drawer  $i$ ) —  $p_i(1 - d_i)$  in total.

This gives the claimed probability,  $p_i(1 - d_i)/(1 - p_i d_i)$ , for the probability that the paper is in drawer  $i$ , given that a search of drawer  $i$  was unsuccessful.

This is an example of *Bayesian search theory*. It has been used in practice, most famously to find the missing nuclear submarine USS Scorpion in 1967. See [http://en.wikipedia.org/wiki/Bayesian\\_search\\_theory](http://en.wikipedia.org/wiki/Bayesian_search_theory) for a little more.

4. The top shelf of the cereal aisle in the grocery store has 6 boxes of nice cereal and 3 boxes of nasty cereal. The middle shelf has 4 nice boxes and 4 nasty boxes. The bottom

shelf has 2 nice boxes and 3 nasty boxes. I choose one box, completely at random, from each shelf (three boxes in all).

- (a) What's the probability that among the three boxes I choose, one is a nice box from the top shelf?

**Solution:**  $6/9 = 2/3$ ; all that matters here is what selection I make among the 9 boxes on the top shelf.

- (b) Given the extra information that I choose exactly two nice boxes in all, what now is the probability that among the three boxes I choose, one is a nice box from the top shelf?

**Solution:** Now what I do on the other shelves matters! Let  $A$  be the event that I choose two nice boxes exactly, and  $B$  the event that I choose a nice box from the top shelf. We want  $\Pr(B|A) = \Pr(B \cap A) / \Pr(A)$ .

For  $\Pr(A)$ : There are three ways to choose exactly two nice boxes in all: from top and middle (probability  $\frac{6.4.3}{9.8.5}$ ), from top and bottom (probability  $\frac{6.4.2}{9.8.5}$ ) and from middle and bottom from (probability  $\frac{3.4.3}{9.8.5}$ ). So  $\Pr(A)$  is the sum of these, or  $.4333\dots$

For  $\Pr(B \cap A)$ : There are two ways to choose exactly two nice boxes in all, including one from the top: from top and middle (probability  $\frac{6.4.3}{9.8.5}$ ), and from top and bottom (probability  $\frac{6.4.2}{9.8.5}$ ). So  $\Pr(B \cap A)$  is the sum of these, or  $.3333\dots$

This makes  $\Pr(B|A) = \Pr(B \cap A) / \Pr(A) \approx .769$  (or exactly  $10/13$ ).

5. I pick three cards at random from a deck of 52 cards, one after the other.

- (a) Compute the probability that the third card is a spade, given that the first two were spades.

**Solution:**

$$\Pr(S_3|S_1 \cap S_2) = \frac{\Pr(S_1 \cap S_2 \cap S_3)}{\Pr(S_1 \cap S_2)} = \frac{\frac{13.12.11}{52.51.50}}{\frac{13.12}{52.51}} = \frac{11}{50}.$$

- (b) Compute the probability that the *first* card is a spade, given that the *last two* were spades.

**Solution:**

$$\Pr(S_1|S_2 \cap S_3) = \frac{\Pr(S_1 \cap S_2 \cap S_3)}{\Pr(S_2 \cap S_3)} = \frac{\frac{13.12.11}{52.51.50}}{\Pr(S_2 \cap S_3)}.$$

What's  $\Pr(S_2 \cap S_3)$ ? One way to find out is to break it up into two probabilities, one each for  $S_1$  also occurring and  $S_1$  not occurring:

$$\Pr(S_2 \cap S_3) = \Pr(S_2 \cap S_3 \cap S_1) + \Pr(S_2 \cap S_3 \cap S_1^c) = \frac{13.12.11}{52.51.50} + \frac{39.13.12}{52.51.50} = \frac{13.12}{52.51}.$$

So  $\Pr(S_2 \cap S_3)$  is exactly the same as  $\Pr(S_1 \cap S_2)$ , and

$$\Pr(S_1|S_2 \cap S_3) = \frac{\frac{13.12.11}{52.51.50}}{\frac{13.12}{52.51}} = \frac{11}{50}.$$

In other words, the probability that the *first* card is a spade, given that the *last* two were spades, is the same as the probability that *last* card is a spade, given that the *first* two were spades.

6. I read the following in a newspaper: “An ectopic pregnancy is twice as likely to develop when the pregnant woman is a smoker, as when she is a non-smoker”. I read in a survey that “32% of all women of childbearing age are smokers”. From these two facts, figure out what percentage of women having ectopic pregnancies are smokers.

**Solution:** The two facts we are given are that  $\Pr(E|S) = 2 \Pr(E|S^c)$ , and  $\Pr(S) = .32$ , from which we want to find  $\Pr(S|E)$ . We know:

$$\Pr(S|E) = \frac{\Pr(S \cap E)}{\Pr(E)} = \frac{\Pr(E|S) \Pr(S)}{\Pr(E|S) \Pr(S) + \Pr(E|S^c) \Pr(S^c)} = \frac{.32 \Pr(E|S)}{.32 \Pr(E|S) + .68 \Pr(E|S^c)}.$$

Substituting in  $\Pr(E|S) = 2 \Pr(E|S^c)$ , we get

$$\Pr(S|E) = \frac{.64 \Pr(E|S^c)}{.64 \Pr(E|S^c) + .68 \Pr(E|S^c)} = \frac{.64}{.64 + .68} \approx .48.$$

7. Andy Murray and Rafael (Rafa) Nadal play a best-of-five-set tennis match (that is, they keep playing sets until one of them has won three; that person is then the winner). Each player has a .5 chance of winning each set, regardless of what happened on previous sets (so, for example, the probability that Andy wins the third set, given that Rafael won the first and Andy the second, is .5). I watch the match, and observe who won each set.

- (a) Write down  $\Omega$ , the sample space for this experiment. Use the following notation: *AARA* stands for the outcome that Andy wins the first set and the second, Rafa the third, and Andy the fourth. **Hint:** your sample space should include exactly 20 outcomes.

**Solution:** The twenty sample points are: *AAA, RAAA, ARAA, AARA, RRAAA, RARAA, RAARA, ARRAA, ARARA, AARRA, RRR, ARRR, RARR, RRAR, AARRR, ARARR, ARRAR, RAARR, RARAR* and *RRAAR*.

- (b) **Use the multiplication rule** to figure out the probability of each of the outcomes: *RRR, AARA, and RAARR*. Use common sense to then write down the probabilities of each of the other outcomes. **Hint:** the sum of your 20 answers should be 1.

**Solution:** Write  $A_i$  (or  $R_i$ ) for the event that Andy (or Rafa) wins the  $i$ th set.

$$\Pr(RRR) = \Pr(R_1 \cap R_2 \cap R_3) = \Pr(R_1) \Pr(R_2|R_1) \Pr(R_3|R_1 \cap R_2) = (.5)(.5)(.5) = .125.$$

Similarly  $\Pr(AARA)$

$$= \Pr(A_1) \Pr(A_2|A_1) \Pr(R_3|A_1 \cap A_2) \Pr(A_4|A_1 \cap A_2 \cap R_3) = (.5)(.5)(.5)(.5) = .0625$$

and  $\Pr(RAARR)$

$$\begin{aligned} &= \Pr(R_1) \Pr(A_2|R_1) \Pr(A_3|R_1 \cap A_2) \Pr(R_4|R_1 \cap A_2 \cap A_3) \Pr(R_5|R_1 \cap A_2 \cap A_3 \cap R_4) \\ &= (.5)(.5)(.5)(.5)(.5) \\ &= .03125. \end{aligned}$$

Common sense suggests that the probability of each sample point is

$$(.5)^{\text{number of sets in that match}} .$$

the two three-set matches each have probability .125, the six four-set matches each have probability .0625 and the twelve five-set matches each have probability .03125 (these numbers do add to 1!)

- (c) Calculate the probability that “Rafa wins the match (gets to three sets before Andy does)”

**Solution:** We add up the probabilities of the ten sample points in which there are three  $R$ 's, and get a total probability of .5 (as we'd expect).

- (d) Calculate the probability that “Rafa has a two sets to one advantage after three sets”

**Solution:** We add up the probabilities of the nine sample points in which there are two  $R$ 's and one  $A$  in the first three slots, and get a total probability of .375.

- (e) Calculate the *conditional* probability that “Rafa wins the match, given that he has a two sets to one advantage after three sets”

**Solution:** This is the ratio: probability of “Rafa wins the match AND has a two sets to one advantage after three sets” to probability of “Rafa has a two sets to one advantage after three sets”. We've calculated the denominator, it's .375. The numerator is .28125 (we add up the probabilities of the nine sample points in which there are two  $R$ 's and one  $A$  in the first three slots, and three  $R$ 's in all). This makes the conditional probability

$$\frac{.28125}{.375} = .75.$$

(Higher than the unconditioned probability. The final answer should make sense: the only way that Andy can win, down two sets to one, is to win both remaining sets, a  $(.5)(.5) = .25$  probability, so Rafa wins with the leftover .75 probability).

8. When I go away for a week, I ask my neighbor to regularly water my plants. I know from experience that she is 90% likely to actually do this. If my plants are regularly watered, they stay healthy with probability .9. If my plants are not regularly watered, they stay healthy with probability .2.

- (a) What is the probability that my plants will be healthy when I return from my trip?

**Solution:** Let  $H$  be the event that my plants are healthy, and  $W$  the event that my neighbor came by to water them. Then

$$\Pr(H) = \Pr(H|W) \Pr(W) + \Pr(H|W^c) \Pr(W^c) = (.9)(.9) + (.2)(.1) = .83.$$

- (b) I come home from my trip, and find that my plants are not healthy. What is the probability that my neighbor *failed* to come by to regularly water them?

**Solution:** By definition of conditional probability,

$$\Pr(W^c|H^c) = \frac{\Pr(W^c \cap H^c)}{\Pr(H^c)} = \frac{\Pr(H^c|W^c) \Pr(W^c)}{\Pr(H^c)} = \frac{(.8)(.1)}{.17} \approx .48.$$

9. Professor Bunsen always starts his Alchemy 231 lecture course with one of the three great alchemical experiments: turning lead into gold (20% of all times that he teaches the course), brewing the elixir of life (40% of the times) and creating the Philosopher's stone (40% of the time). When he tries to turn lead into gold, the result always ends with a explosion; when he brews the elixir of life, there is a 50% chance of an explosion, and when he creates the Philosopher's stone, 8 times out of 10 there is an explosion. Dean Crawford wants to see which experiment Professor Bunsen will do this year, but he arrives late. If he see the lecture-hall filled with post-explosion smoke, what should he conclude is the probability that he has just missed a demonstration of brewing the elixir of life?

**Solution:** With the names for the events hopefully self-explanatory,

$$\begin{aligned} \Pr(EL|SMOKE) &= \frac{\Pr(EL \cap SMOKE)}{\Pr(SMOKE)} \\ &= \frac{\Pr(SMOKE|EL) \Pr(EL)}{\Pr(SMOKE|EL) \Pr(EL) + \Pr(SMOKE|LG) \Pr(LG) + \Pr(SMOKE|PS)} \\ &= \frac{(.5)(.4)}{(.5)(.4) + (1)(.2) + (.8)(.4)} \\ &\approx .278. \end{aligned}$$

10. 98% of the time in which a person is telling the truth, a lie detector correctly detects this. 95% of the time in which a person is lying, the lie detector correctly detects this. An operator knows from experience that when suspects are being interrogated, they lie two thirds of the time. Based on this information, what would you say is the probability that a statement detected as a lie is in fact a lie? (Let  $D$  be the event that a statement is detected as a lie,  $T$  the event that a statement is not a lie and  $L$  the event that a statement is a lie).

**Solution:**

$$\begin{aligned}\Pr(L|D) &= \frac{\Pr(L \cap D)}{\Pr(D)} \\ &= \frac{\Pr(D|L) \Pr(L)}{\Pr(D|L) \Pr(L) + \Pr(D|T) \Pr(T)} \\ &= \frac{(.95)(.66)}{(.95)(.66) + (.02)(.33)} \\ &\approx .9896.\end{aligned}$$

11. Experience suggests that when a company sends a survey out, the response rate among under-25's is 30% and the response rate among over-45's is 55%. A survey is sent out to a bunch of people; 30% of the people are under-25's, 30% are over-45's, and the rest are 26-to-44's. There is a 40% response rate. Assuming that the response rates for this particular survey are typical, what do you think is the usual response rate from 26-to-44's?

**Solution:** Let  $p$  be the usual response rate from 26-to-44's. From the law of total probability and the data given we expect

$$.4 = (.3)(.3) + (.4)p + (.3)(.55)$$

so  $p = .3625$ .

12. If  $\Pr(A) > 0$ , show that  $\Pr(A \cap B|A) \geq \Pr(A \cap B|A \cup B)$ .

**Solution:** The left-hand side is

$$\Pr(A \cap B|A) = \frac{\Pr((A \cap B) \cap A)}{\Pr(A)} = \frac{\Pr(A \cap B)}{\Pr(A)}.$$

The right-hand side is

$$\Pr(A \cap B|A \cup B) = \frac{\Pr((A \cap B) \cap (A \cup B))}{\Pr(A \cup B)} = \frac{\Pr(A \cap B)}{\Pr(A \cup B)}.$$

So we want to show

$$\frac{\Pr(A \cap B)}{\Pr(A)} \geq \frac{\Pr(A \cap B)}{\Pr(A \cup B)}$$

which is the same as

$$\Pr(A \cup B) \geq \Pr(A),$$

which is true since  $A \subseteq A \cup B$ .

13. For events  $A$  and  $B$  with  $\Pr(A) > 0$ ,  $\Pr(B) > 0$ , we say that  $A$  is *positively correlated* with  $B$  if  $\Pr(A|B) > \Pr(A)$  (in other words,  $B$  occurring *increases* the chance that  $A$  occurs). If  $\Pr(A|B) < \Pr(A)$  then  $A$  is *negatively correlated* with  $B$ . (If  $\Pr(A|B) = \Pr(A)$  then  $A$  is *not correlated* with  $B$ .)

- (a) Show that if  $A$  is positively correlated with  $B$ , then  $B$  is positively correlated with  $A$ ; in other words, show that if  $\Pr(A|B) > \Pr(A)$  then  $\Pr(B|A) > \Pr(B)$ . (The same holds for negatively correlated and uncorrelated; so we can from now on just say “ $A$  and  $B$  are positively correlated”, etc.)

**Solution:**

$$\begin{aligned} \Pr(A|B) > \Pr(A) &\iff \frac{\Pr(A \cap B)}{\Pr(B)} > \Pr(A) \\ &\iff \Pr(A \cap B) > \Pr(A) \Pr(B) \\ &\iff \frac{\Pr(A \cap B)}{\Pr(A)} > \Pr(B) \\ &\iff \Pr(A|B) > \Pr(A). \end{aligned}$$

- (b) Looking back at the earlier question about Andy Murray and Rafael Nadal, use your answers there to figure out whether the events “Rafa wins the match” and “Rafa has a two sets to one advantage after three sets” are positively correlated, negatively correlated, or uncorrelated.

**Solution:** We calculated that the probability that Rafa wins, given that he has a two sets to one advantage after three sets, is .75. This is greater than the unconditioned probability that he wins, which is .5. So these events are positively correlated.

- (c) I roll two dice and note the numbers that come up. Are the events “The sum of the numbers coming up is even” and “The two numbers are within one of each other” positively correlated, negatively correlated, or uncorrelated?

**Solution:** Probability of sum being even is  $18/36$ . Probability of numbers being within one of other is  $16/36$ . Probability of sum being even AND numbers being within one of other is  $6/36$ . Since  $6/36 < (18/36)(16/36)$ , the events are negatively correlated.

14. Chapter 1, problem 22. (Let  $A_i$  be the event that the ball drawn from the  $i$ th jar is white. Calculate  $\Pr(A_1)$ , and then use your knowledge of  $\Pr(A_1)$  to calculate that  $\Pr(A_2) = \Pr(A_1)$ . Then show that if you know  $\Pr(A_i) = \Pr(A_1)$ , you also have that  $\Pr(A_{i+1}) = \Pr(A_1)$ . This is enough (by induction) to show that  $\Pr(A_k) = \Pr(A_1)$  for every possible  $k$ .)

**Solution:**

$$\Pr(A_1) = \frac{m}{m+n}.$$

To calculate  $\Pr(A_2)$ , we use the law of total probability. The sample space divides into two pieces: sample points in which  $A_1$  occurred (so jar 2 has  $m+1$  white balls and  $n$



black), and sample points in which  $A_1$  did not occur (so jar 2 has  $m$  white balls and  $n + 1$  black). By the law of total probability,

$$\begin{aligned} \Pr(A_2) &= \Pr(A_2|A_1) \Pr(A_1) + \Pr(A_2|A_1^c) \Pr(A_1^c) \\ &= \left(\frac{m+1}{m+n+1}\right) \left(\frac{m}{m+n}\right) + \left(\frac{m}{m+n+1}\right) \left(\frac{n}{m+n}\right) \\ &= \frac{m}{m+n} \\ &= \Pr(A_1). \end{aligned}$$

To calculate  $\Pr(A_{i+1})$ , we again use the law of total probability. The sample space divides into two pieces: sample points in which  $A_i$  occurred (so jar  $i + 1$  has  $m + 1$  white balls and  $n$  black), and sample points in which  $A_i$  did not occur (so jar  $i + 1$  has  $m$  white balls and  $n + 1$  black). By the law of total probability,

$$\begin{aligned} \Pr(A_{i+1}) &= \Pr(A_{i+1}|A_i) \Pr(A_i) + \Pr(A_{i+1}|A_i^c) \Pr(A_i^c) \\ &= \left(\frac{m+1}{m+n+1}\right) \Pr(A_i) + \left(\frac{m}{m+n+1}\right) (1 - \Pr(A_i)). \end{aligned}$$

If we assume (by induction) that  $\Pr(A_i) = \Pr(A_1) = m/(m+n)$ , then exactly as before this gives

$$\Pr(A_{i+1}) = \left(\frac{m+1}{m+n+1}\right) \left(\frac{m}{m+n}\right) + \left(\frac{m}{m+n+1}\right) \left(\frac{n}{m+n}\right) = \frac{m}{m+n} = \Pr(A_i).$$

So, by induction,  $\Pr(A_i) = \Pr(A_1) = m/(m+n)$  for all  $i$  (including in particular  $i = k$ ).