**Answer Key to DAT 520: DA Problem Set 3: Cumulative and Conditional Probability**



The last problem set exposed you to dependent trials that caused the probability to shift slightly with each choice of sock from the drawer. Kind of sneaky, eh? This problem set gets you a little further into conditional probability with a direct look at Bayes’ Theorem. Since this is a course in Decision Analysis, not a course in probability, we are going to introduce these concepts, use them once or twice by hand, but then let R and Excel do the heavy lifting for us the rest of the time. This problem set, however, is hand calculations. All you should need is a calculator and your wits.

*Example 1*: You are flipping a fair coin 4 times total, but in groups of 2. In order to flip the coin for flips 3 and 4, you have to get heads two times in a row for flips 1 and 2.

Questions: a) What is the probability of flipping heads 4 times in a row, *generally*?

b) What is the *total probability* of getting 4 heads, the way this problem is set up?

c) What is the probability that you will get two heads on flips 3 and 4, *conditional* on getting two heads on flips 1 and 2?

Answer: **Define success**. Success here is flipping a heads, which is p=50%.

a) Flipping 4 heads in a row is “4 choose 4” with p=.5 So, using the binomial equation:



We calculate this as **.0625** chance of flipping four heads on four flips, *generally*.

b) There are two identical events of two coin flips each. So, each one is “2 choose 2” with p=.5 ...use the **binomial equation to calculate cumulative probability**:



We calculate this as 25% chance of flipping two heads on flips 1 and 2. Same thing for flips 3 and 4. Let’s call the first event p(F1&2) and the second one p(F3&4).

Now, **build the table**. Notice that all the % contributions have to add up to 1.

Table of contributions to total probability:

P %contrib

p(F1&2) .25 .50

p(F3&4) .25 .50

Use the **Law of Total Probability**:

probability(item\_1) x (item\_1\_%conribution) + probability(item\_2) x (item\_2\_%conribution) + … etc.

Total Probability for 2 heads then 2 heads is (.50 x .25) + (.50 x .25) = **25%.**

c) Probability of two heads on p(F3&4) conditional on two heads for p(F1&2).

Bayes’ Theorem: 

That’s not very useful in this form. Here, we are looking for the probability of the second item, p(F3&4), given the first, p(F1&2). The vertical bar means “given”.



Which still isn’t very helpful. We didn’t calculate it this way. We didn’t calculate p(F1&2|F3&4). If we had, we would know the answer is just 1 minus the result. But we DID calculate each part’s relation to the whole. So **use the form of Bayes’ Theorem that uses the Law of Total Probability as the denominator**:



Which is p(F3&4)’s *proportion* to the total probability. From the table we made, we already know all of those pieces on the right-hand side:



**Interpret.** The probability that I will flip 2 heads in a row on the condition of previously flipping 2 heads in a row is 50%. Bear in mind that the raw probability of flipping two heads in a row in independent trials is 25% and flipping four heads in a row in independent trials is 6.25%.

***Recap of the method:***

1. Define success of an individual trial.

2. Use the binomial equation to calculate the cumulative probability of each part of the system.

3. Make a table with each part’s p and its % contribution to the whole. The %s must add to 1.

4. Calculate the p(Total) using the Law of Total Probability.

5. Use the Total Probability form of Bayes’ Theorem to calculate the *Conditional Probability* for the item of interest.

6. Interpret.

1. Answer: 10 choose 1, with p=0.25: 18.75%
2. Answer: 10 choose 7, with p=0.25: 0.31%
3. Answer: p success=33.3%, fail = 66.7%: 1.6% chance. 5x better, but still small

Exercises:

1. You are taking a 10 question multiple choice test. If each question has 4 choices and you guess on each question, what is the probability of getting 1 question correct? [Hint: This is just a binomial.]

[Answer: 10 choose 1, with p=0.25: 18.75%

*n* = 10

*k* = 1

*n – k* = 9

*p* = 0.25 = probability of guessing the correct answer on a question

*q* = 0.75 = probability of guessing the wrong answer on a question

]

2. What is the probability of getting 7 questions correct? [Hint: So is this one.]

*[*Answer: 10 choose 7, with p=0.25: 0.31%

*n* = 10

*k* = 7

*n – k* = 3

*p* = 0.25 = probability of guessing the correct answer on a question

*q* = 0.75 = probability of guessing the wrong answer on a question

]

3. What are your chances of getting 7 right if you can reliably eliminate one possible answer from each question? [Hint: This one, too.]

[Answer: p success=33.3%, fail = 66.7%: 1.6% chance. 5x better, but still small.]

4. Let’s say, instead, that the test is an adaptive test: you get to answer more questions based on your previous success.

This annoying test is structured like this: first you have to answer 3 questions and if you are correct on two of them, you get to answer three more questions and if two of *those* are correct, then you get 3 final questions, of which you need to get at least 2 right to pass the whole test.

The first test T1 has 3 multiple choice questions with 4 possible answers each (p=0.25 per question.)

The second test T2 has 3 multiple choice questions with 3 possible answers each (p=0.33 per question.)

The final test T3 has 3 questions that are true/false (p=0.50 each question.)

These questions are in a language you have never seen: a mixture of Navaho, Swahili, Klingon and Esperanto. So you have to guess on all of the questions and there are no contextual clues to eliminate any answers. This is the first one:

*'Arlogh Qoylu'pu'?*

Moja: Yel kholgo eeah.

Mbili: Floroj kreskas ĉirkaŭ mia domo. Pe'el!

Tatu: La sandviĉo estos manĝota'mo'tlhIngan maH!

Nne: 'Adeez'æ`q eeah.

(The professor sits at the front of class with a giant, sadistic grin while the students throw wads of paper at his head.)

Using the Binomial Probability Rule, the Law of Total Probability and Bayes’ Theorem,

a) What is the probability of getting 2 right on each sub-exam? (T1, T2 and T3, separately.)

b) What are your overall chances of passing the entire exam?

c) What are your chances of passing T3, if you first pass T1 and T2?

Hint: Structure your analysis. Figure out the component probabilities: p(passing test 1), p(passing test 2), p(passing test 3). Then make a table of their proportional contributions of probability to the whole. Then calculate the total probability: p(Total). Finally, use Bayes’ Theorem to calculate the probability of passing test 3 conditional on passing tests 1 & 2.

[Answer:

Part 1. Probability of getting 2 right on each test of 3 questions, purely by chance. This is the familiar binomial probability for independent trials:







Additionally, the probability of getting 3 right on each test is:

p(P3|T1) = 0.015625

p(P3|T2) = 0.035937

p(P3|T3) = 0.125

So the probability of 2 or 3 right is

p(P2orP3|T1) = 0.15625

p(P2orP3|T2)=0.254826

p(P2orP3|T3)=0.5

Part 2. Then structure the table of each test’s contribution to the whole:

Test %contribution p(of passing it)

T1 p(%T1)=.33 p(Passing|T1)= 0.15625

T2 p(%T2)=.33 p(Passing|T2)= 0.254826

T3 p(%T3)=.33 p(Passing|T3)= 0.5

The vertical bar | means “given”. Say it aloud: p(Passing|T1). “Probability of passing, given test 1.” Which means “the probability of passing test 1,” which you just calculated using the binomial probability.

Part 3. The law of total probability. This is the overall probability of passing all exams, as if they were independent. It’s just the probabilities of passing each test multiplied by its individual contribution to the whole system, in our case each one is worth ⅓ of the total:

p(Total) = (0.33 x .15625) + (0.33 x 0.254826) + (0.33 x 0.5) = 0.30065508

You have a 30% probability of passing the entire exam if presented with all of its sections all at once..

Part 4. Finally calculate the probability of passing T3 conditional on previously passing T1 and T2, since passing the whole test is conditional upon passing all of its parts in order. We need to use Bayes’ theorem (another dialect of Klingon) to stack those probabilities against the overall event (total probability) of passing.



= 0.5 x 0.33 = .54880164

0.30065508

*55%*

This is actually amazing. What is it saying? It’s saying that if you pass the first two tests, your chances of passing the last one (which are True/False) are better than 50%! How in the heck? That is the beauty of Bayesian Stats. It can be counter-intuitive. I’m telling you… it can get very deep, especially when we move into the more complex forms of modeling.

5. Let’s say that you know just enough of these obscure languages to translate the first question in T1:

What time is it? (Klingon)

1 (Swahili): From dawn to setting sun. (Navajo)

2 (Swahili): Flowers grow around my house (Esperanto) so all of you may come in. (Klingon)

3 (Swahili): The sandwich will be eaten (Esperanto) because we are Klingons! (Klingon)

4 (Swahili): It’s mid-afternoon. (Navajo) [correct answer]

Now the probability of passing T1 has changed because you only have to guess correctly on one of the two remaining questions in the first section.

a) What is the new probability for T1?

b) Now what is the overall probability of passing the entire test?

c) And what is the probability of passing section T3, given that you’ve already passed sections T1 and T2?

d) The kicker: How do you explain the difference between 4c and 5c? Can you relate this to a larger context about conditional probability and making decisions?

[Answer:

a) T1’s new probability is 0.375 (2 choose 1, with p=0.25) + 0.0625 (2 choose 2, with p=0.25) = 0.4375

b) New total probability is 0.3989.

c) The revised probability of passing T3 conditional on previously passing T1 and T2 is

0.5 \* 0.33 / 0.3989 = 0.4052

d) Why has it gone down? After all, we eliminated a question in section 1, so shouldn’t it be easier to pass the test?

Keep in mind that this is *conditional probability*. They are proportions to a whole system that must add up to 1. If one part of a system has been re-valued *higher*, then the other parts of the system will have to be re-valued *lower*. Now in this question, Test 3’s contribution to passing the entire exam has been reduced. Which is still greater than if the entire test were merely a summation of three individual trials. But it’s not. Each section of the test depends upon the previous sections.

This is a lot to think about. But what we’ve done here is examine in a mathematical way *everything you need to know about Markov Models and probability revision*. Good work, everyone! ]

---

Google Latex formulae:

{p(F3&4|F1&2) =} \frac{p(F1&2)p(%contribF1&2)}{p(Total)}

=\frac{.25\cdot .50}{.25}=50%

{p(Passing|T\_{1})}=}{3 \choose 2} (0.25)^{2}(0.75)^{1} {= 0.140625}

{p(Passing|T\_{2})}=}{3 \choose 2} (0.33)^{2}(0.67)^{1} {= 0.218889}

{p(Passing|T\_{3})}=}{3 \choose 2} (0.50)^{2}(0.50)^{1} {= 0.375}

{p(T\_{3}|Passing) =} \frac{p(Passing|T\_{3})p(%T\_{3})}{p(Passing)}

Bayes’ Theorem:

{p(A|B) =} \frac{p(B|A)p(A)}{p(B)}

{p(F3&4|F1&2) =} \frac{p(F1&2|F3&4)p(%F3&4)}{p(F1&2)}

=\frac{.25\cdot .50}{.25}=50%

{p(F3&4|F1&2) =} \frac{p(F1&2|F3&4)p(F3&4)}{p(F1&2)}

{p(F3&4|F1&2) =} \frac{p(F1&2) \cdot p(%contribF1&2)}{p(Total)}