MIT Introduction to Statistics 18.05 Problem Set 2

John Hancock

March 4, 2017

Contents

1 References and License

We are answering questions in the material from MIT OpenCourseWare course 18.05, Introduction to Probability and Statistics.
Please see the references section for detailed citation information.
The material for the course is licensed under the terms at `http://ocw.mit.edu/terms`.
We are answering the questions that Orloff and Bloom ask in [6].
We use documentation in [2] to write LaTeXsource code for this document.

2 'Boy or girl' paradox

In order to write this solution, we rely on the answer to this problem in [7], and the treatment of the 'Boy or girl,' paradox in [1].
For these questions on the 'Boy or girl paradox we deal with events B, "the child is a boy," and G, "the child is a girl."

We assume B, and G have the same properties as the B and G events Orloff and Bloom analize in example 9 of [9]. These properties are that B, and G are independent, and they have probability $\frac{1}{2}$.
We use these properties to define 4 more events, BB, BG, GB, and GG. These events are: "the younger child is a boy, and the older child is a boy," "the younger child is a boy, and the older child is a girl," "the younger child is a girl, and the older child is a boy," "the younger child is a girl, and the older child is a girl," respectively. We use the properties of B, and G, of example 9 to compute that the probabilities of BB, BG, GB, GG, $P(BB)$, $P(BG)$, $P(GB)$, $P(GG)$, are all equal to $\frac{1}{4}$.

2.1 Probability of girls

The question Orloff and bloom quote is, "Mr. Jones has two children. The older child is a girl. What is the probability that both children are girls?"
We can restate the question above as, "Given event BG or event GG, what is the probability GG?"
We use the definition of conditional probability. We also use the law of total probability to comupte $P(GG \cup BG)$.
Therefore we write the equation:

$$P(GG \mid GG \cup BG) = \frac{P(GG \cap (GG \cup BG))}{P(GG \cup BG)} \tag{1}$$

$$\frac{P(GG \cap (GG \cup BG))}{P(GG \cup BG)} = \frac{P(GG)}{P(GG \cup BG)} \tag{2}$$

$$\frac{P(GG)}{P(GG \cup BG)} = \frac{\frac{1}{4}}{\frac{1}{2}} \tag{3}$$

$$\frac{\frac{1}{4}}{\frac{1}{2}} = \left(\frac{1}{4}\right)\left(\frac{2}{1}\right) = \frac{1}{2} \tag{4}$$

Therefore if Mr. Jones' older child is a girl, there is a probabiility of $\frac{1}{2}$ that the younger child is also a girl.

2.2 Probability of boys

In this section, Orloff and Bloom quote another question for us to answer here.
The question is, "Mr. Smith has two children. At least one of them is a boy. What is the probability that both children are boys?"
We use the definitions from the previous section for events, BB, BG, GB and GG. We use the probabilities we found in the first section for these events as well.

The author of this question is giving us that three possible events have occured: BB, BG, or GB. Furthermore the question asks for the conditional probability of BB.
We use the definition of conditional probability, and the law of total probability to compute:

$$P(BB \mid BB \cup BG \cup GB) = \frac{P(BB \cap (BB \cup BG \cup GB))}{P(BB \cup BG \cup GB)} \tag{5}$$

$$\frac{P(BB \cap (BB \cup BG \cup GB))}{P(BB \cup BG \cup GB)} = \frac{P(BB)}{P(BB \cup BG \cup GB)} \tag{6}$$

$$\frac{P(BB)}{P(BB \cup BG \cup GB)} = \frac{\frac{1}{4}}{\frac{1}{4}+\frac{1}{4}+\frac{1}{4}} = \left(\frac{1}{4}\right)\left(\frac{4}{3}\right) = \frac{1}{3} \tag{7}$$

If at least one of Mr. Smith's children is a boy, then there is a probability of $\frac{1}{3}$ that both children are boys.

3 The blue taxi

In order to solve this problem we will write a confusion matrix [3]. This is the term we found for the kind of table Orloff and Bloom write in [8] that organizes false positive rates, false negative rates, etc. Into a table.
We define the following sets:

- $D+$, "The car is blue."
- $D-$, "The car is green."
- $T+$, "The witness reports seing a blue car."
- $T-$, "The witness reports seing a green car."

Orloff and Bloom give us the following probabilities:

- $P(D+) = 0.01$
- $P(D-) = 0.99$
- $P(T+ \mid D+) = 0.99$
- $P(T+ \mid D-) = 0.02$

In order to make our case, we need to know $P(D+ \mid T+)$. That is the probability that, given a blue car, the witness saw a blue car.
This table summarizes the information we know. Note the small ratio of blue taxis to all taxis: $\frac{1}{100}$.

	Green	Blue
Sees Blue	$P(T+ \mid D-) = 0.02$	$P(T+ \mid D+) = 0.99$
Total	$P(D-) = 0.99$	$P(D+) = 0.01$

We apply Bayes' theorem [8] to $P(T+ \mid D+)$ in order to compute $P(D+ \mid T+)$.

$$P(D+ \mid T+) = \frac{P(T+ \mid D+)\,P(D+)}{P(T+)} \tag{8}$$

We use the Law of total probability to rewrite the denominator of the fraction on the righthand side of 8

$$\frac{P(T+ \mid D+)\,P(D+)}{P(T+)} = \frac{P(T+ \mid D+)\,P(D+)}{P(T+\cap D+) + P(T+\cap D-)} \tag{9}$$

We now use the definition of conditional probabilty to rewrite the probabilities in the denomenator of the equation in the right hand side of the equation above:

$$\frac{P(T+ \mid D+)\,P(D+)}{P(T+)} = \frac{P(T+ \mid D+)\,P(D+)}{P(T+ \mid D+)\,P(D+) + P(T+ \mid D-)\,P(D-)} \tag{10}$$

The terms of the righthand side of the equation above are all in our table, so we now have a way to compute $P(D+ \mid T+)$:

$$\frac{P(T+ \mid D+)\,P(D+)}{P(T+ \mid D+)\,P(D+) + P(T+ \mid D-)\,P(D-)} = \frac{0.99 \times 0.01}{0.99 \times 0.01 + 0.02 \times 0.99} \tag{11}$$

Now we simplify the right hand side of the euqation above to arrive at a value for $P(D+ \mid T+)$.

$$\frac{0.99 \times 0.01}{(0.01 + 0.02) \times 0.99} = \frac{0.99 \times 0.01}{0.03 \times 0.99} = \frac{1}{3} \tag{12}$$

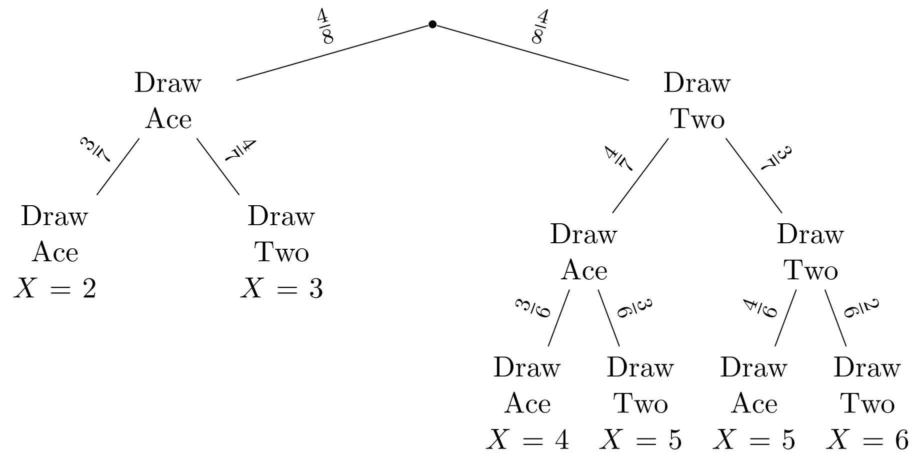
Therefore there is a $\frac{1}{3}$ that given a blue taxi, the witness sees a blue taxi. This is a less than 50% chance that the witness actually saw a blue taxi. Hence we have a reasonable doubt that the witness saw a blue taxi.

4 Trees of cards

In this section we answer Orloff and Bloom's question in [6] about the expected value of a random variable.
The random variable is the value of the sum of cards we draw from a hat.
We refer to a card of rank one as an ace, and a card of rank two as a two.
There are four aces, and four two's in the hat.
Rules govern the way we draw the cards from the hat. The rules are:

- If we draw an ace first, then we draw one more card.
- If we draw a two first, then we draw two more cards.

We assign one to the value of the ace, and two to the value of the two.
We assign the random variable X the value of the sum of the values of the cards we draw.



We supplement terms we use when we write about trees with definitions of leaf, root from [5]. We also use definitions of vertex, node, edge, and path in [4] when we write about trees.
Recall from [8] that the fractions on the edges of the probability trees like the tree above are probabilities, and that the nodes of the tree are events. Nodes that are not directly connected to the root of the tree are unions of events. Furthermore, we learn in [8] that the probabilities on the edges of probability trees that are not connected to the root of the tree are conditional probabilities, where the condition given is the event that the node connected to the edge that is closer to the root of the tree represents.
[8] also allows us to multiply probabilities on paths from the root of the probability tree to the leaves of a probability tree to compute the probability of the events that are the leaf nodes of the tree.
Now we are armed with the facts that enable us to compute the expected value $E\left(X\right)$.
We use the definition of expected value in [10]:

$$\begin{aligned} E\left(X\right) = 2\left(\frac{4}{8}\right)\left(\frac{3}{7}\right) + 3\left(\frac{4}{8}\right)\left(\frac{4}{7}\right) + 4\left(\frac{4}{8}\right)\left(\frac{4}{7}\right)\left(\frac{3}{6}\right) \\ +5\left(\frac{4}{8}\right)\left(\frac{4}{7}\right)\left(\frac{3}{6}\right) + 5\left(\frac{4}{8}\right)\left(\frac{3}{7}\right)\left(\frac{4}{6}\right) + 6\left(\frac{4}{8}\right)\left(\frac{3}{7}\right)\left(\frac{2}{6}\right) \end{aligned} \tag{13}$$

We write a statement in the R programming language to calculate the value of the right hand side of equation 13:

```
> 2*(4/8)*(3/7) + 3*(4/8)*(4/7) + 4*(4/8)*(4/7)*(3/6)
+ 5*(4/8)*(4/7)*(3/6) + 5*(4/8)*(3/7)*(4/6) + 6*(4/8)*(3/7)*(2/6)
[1] 3.714286
```

Note: we inserted a line break in the R statement above in order to fit it onto the page; the reader will need to remove the line break in order to reproduce the result.
Therefore $E(X) \approx 3.714$.

5 Dice

In this section, we answer problem for that Orloff and Bloom pose in [6].

5.1 Probability Mass Function

The first query Orloff and Bloom make is, "What is the pmf of S."
We find the definition of pmf in [10]. Orloff and Bloom give the definition of the random variable S in [6]. Please see [9] for the definition of a random variable.
We use a table to give the pmf of S.

k	4	6	8
pmf $P(S = k)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$

5.2 Apply Bayes' Rule

In this part of the question, Orloff and Bloom ask us to find the conditional probabilities $P(S = k \mid R = 3)$, for $k = 4$, $k = 6$, and $k = 8$.
Orloff and Bloom also give a terminology note for this problem, writing, "You are computing the pmf of S given R = 3." [6]
We will apply Bayes rule to compute $P(S = k \mid R = 3)$ because it is easy to compute $P(R = 3 \mid S = k)$.
$P(R = 3 \mid S = k)$ is the probability that we roll a 3 with a k-sided die.
We assume the probabilities of the values we roll with any of the dice Orloff and Bloom define in this problem are all equally likely.
We write a table to list the values of $P(R = 3 \mid S = k)$, for $k = 4$, $k = 6$, and $k = 8$.

k	4	6	8
$p(R = 3 \mid S = k)$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{8}$

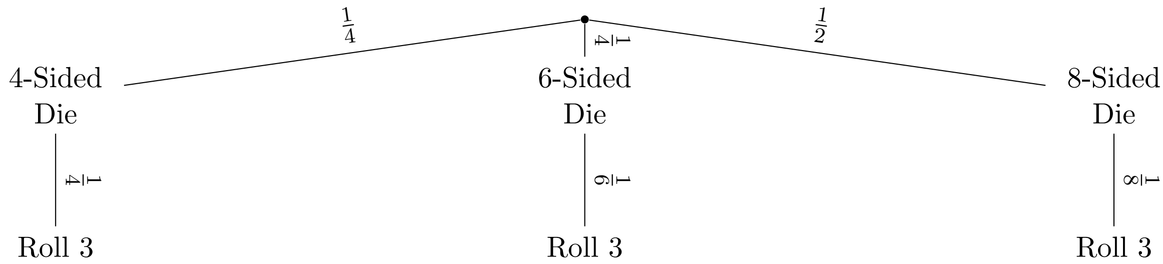
Let us apply Bayes theorem [8] to $P(S = k \mid R = 3)$:

$$P(S = k \mid R = 3) = \frac{P(R = 3 \mid S = k)\, P(S = k)}{P(R = 3)} \tag{14}$$

Equation 14 gives us a formula that we can apply to the second row of the elements of table above to compute the values of $P\left(S=k \mid R=3\right)$, for $k=4$, $k=6$, and $k=8$. We give these formulas in the table below:

k	4	6	8
$p\left(R=3 \mid S=k\right)$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{8}$
$p\left(S=k \mid R=3\right)$	$\frac{P(R=3 \mid S=4)P(S=4)}{P(R=3)}$	$\frac{P(R=3 \mid S=6)P(S=6)}{P(R=3)}$	$\frac{P(R=3 \mid S=8)P(S=8)}{P(R=3)}$

We are almost ready to compute $P\left(S=k \mid R=3\right)$, for $k=4$, $k=6$, and $k=8$, but we lack a value for $p\left(R=3\right)$.
We will draw a probability tree, and use the law of total probability to compute $P\left(R=3\right)$.



The probabilities on the edges connected to the root of the probablity tree above are from the probability mass function for S that we gave in the previous section. These are the probabilities of the event of selecting a 4, 6, or 8 sided die.
The probabilities on the edges connected to the leaf nodes are the conditional probabilities we computed for $P\left(R=3 \mid S=k\right)$ in the table above.
Please note that we use the same definitions and operations on the elements of probability trees here that we use in section 4.
The probability of the event of any one of the leaf nodes in the tree above is the product of the probabilities that label the edges on the path to one of the leaf nodes.
Hence, we apply the law of total probability to compute the total probability of the events that are the leaf nodes of the tree above:

$$P\left(R=3\right)=\left(\frac{1}{4}\right)\left(\frac{1}{4}\right)+\left(\frac{1}{4}\right)\left(\frac{1}{6}\right)+\left(\frac{1}{2}\right)\left(\frac{1}{8}\right) \tag{15}$$

In order to complete the calculation, we use the common denomenator of 48:

$$\left(\frac{1}{4}\right)\left(\frac{1}{4}\right)+\left(\frac{1}{4}\right)\left(\frac{1}{6}\right)+\left(\frac{1}{2}\right)\left(\frac{1}{8}\right)=\left(\frac{3}{48}\right)+\left(\frac{2}{48}\right)+\left(\frac{3}{48}\right) \tag{16}$$

We add the fractions, and reduce to lowest terms:

$$\left(\frac{3}{48}\right) + \left(\frac{2}{48}\right) + \left(\frac{3}{48}\right) = \frac{8}{48} = \frac{1}{6} \tag{17}$$

Therefore $P\left(R=3\right) = \frac{1}{6}$

Now we can compute the values of $P\left(S=k \mid R=3\right)$, for $k=4$, $k=6$, and $k=8$. We give these values in the table below. Please see our comments on the probability tree above for information on how we know $P\left(S=k\right)$ for $k=4$, $k=6$, and $k=8$.

k	4	6	8
$p\left(R=3 \mid S=k\right)$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{8}$
$p\left(S=k \mid R=3\right)$	$\frac{P(R=3\mid S=4)P(S=4)}{P(R=3)}$	$\frac{P(R=3\mid S=6)P(S=6)}{P(R=3)}$	$\frac{P(R=3\mid S=8)P(S=8)}{P(R=3)}$
$p\left(S=k \mid R=3\right)$	$\frac{\left(\frac{1}{4}\right)\left(\frac{1}{4}\right)}{\frac{1}{6}}$	$\frac{\left(\frac{1}{6}\right)\left(\frac{1}{4}\right)}{\frac{1}{6}}$	$\frac{\left(\frac{1}{8}\right)\left(\frac{1}{2}\right)}{\frac{1}{6}}$
simplify above	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{3}{8}$

The last row of the table above gives the values for $P\left(S=k \mid R=3\right)$, for $k=4$, $k=6$, and $k=8$ that Orloff and Bloom require as a solution to this problem.

In this problem, Orloff and Bloom ask a follow up question, "Which die is most likely if $R=3$?". The largest value in the last row in the table above is $\frac{3}{8}$, so the 4-sided and the 8-sided die are equally likely if $R=3$.

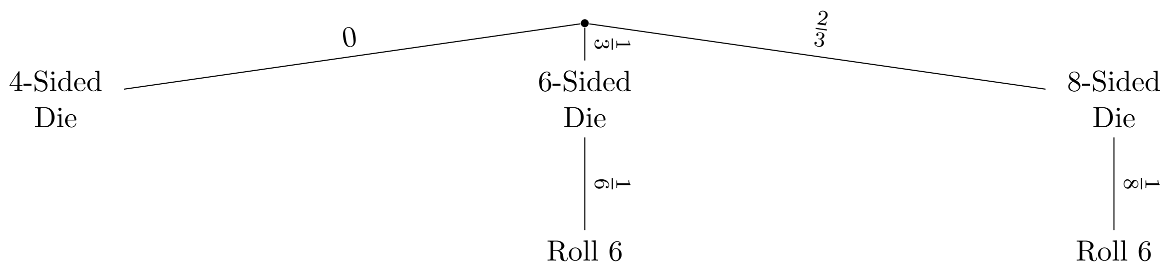
5.3 Rolling a six

This section is similar to the previous section, except that here Orloff and Bloom are asking us, "Which die is most likely if R = 6?" [6]

This question is equivalent to: what is the maximum value of $P\left(S=k \mid R=6\right)$ for values of $k=4$, $k=6$, and $k=8$.

We use the probability mass function $P\left(S=k\right)$ for the random variable S from the previous section.

We draw a probability tree for $P\left(R=6\right)$



We use the rules governing probability trees from [8] to compute $P\left(R=6\right)$:

$$P\left(R=6\right)=\left(\frac{1}{3}\right)\left(\frac{1}{6}\right)+\left(\frac{2}{3}\right)\left(\frac{1}{8}\right) \tag{18}$$

We can do some artihmetic to simplify the right hand side of equation 19:

$$\left(\frac{1}{3}\right)\left(\frac{1}{6}\right)+\left(\frac{2}{3}\right)\left(\frac{1}{8}\right)=\frac{1}{3}\left(\frac{1}{6}+\frac{2}{8}\right)=\frac{1}{3}\left(\frac{4}{24}+\frac{6}{24}\right)=\frac{10}{72}=\frac{5}{36}. \tag{19}$$

Now we have all the probabilities necessary to fill out a table similar to the last table in the previous section to determine the conditional probabilities $P\left(S=k \mid R=6\right)$ for values of $k=4$, $k=6$, and $k=8$:

k	4	6	8
$p\left(R=6 \mid S=k\right)$	0	$\frac{1}{6}$	$\frac{1}{8}$
$p\left(S=k \mid R=6\right)$	Bayes' Theorem Does Not Apply	$\frac{P(R=6\mid S=6)P(S=6)}{P(R=6)}$	$\frac{P(R=6\mid S=8)P(S=8)}{P(R=6)}$
$p\left(S=k \mid R=6\right)$	Bayes' Theorem Does Not Apply'	$\frac{\left(\frac{1}{6}\right)\left(\frac{1}{3}\right)}{\frac{5}{36}}$	$\frac{\left(\frac{1}{8}\right)\left(\frac{2}{3}\right)}{\frac{5}{36}}$
simplify above	Bayes Theorem Does Not Apply	$\frac{2}{5}$	$\frac{3}{5}$

In the table above, we see that the maximum value of $P\left(R=6 \mid S=k\right)$ for values of k, $k=4$, $k=6$, and $k=8$, is $P\left(R=6 \mid S=8\right)=\frac{3}{5}$.

References

[1] 113.161.72.37 et al. *Boy or Girl paradox*. Available at `https://en.wikipedia.org/w/index.php?title=Boy_or_Girl_paradox&oldid=766674814` (Spring 2014).

[2] 84.173.248.72 et al. *LaTeX/Special Characters*. Available at `https://en.wikibooks.org/w/index.php?title=LaTeX/Special_Characters&stable=1` (2016/12/28).

[3] Howard Hamilton. *Confusion Matrix*. Available at `http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion_matrix/confusion_matrix.html` (2012/6/8).

[4] Milos Hauskrecht. *Graphs*. Available at `https://people.cs.pitt.edu/~milos/courses/cs441/lectures/Class25.pdf` (2017/3/4).

[5] Pětr Hlineý. *Basics of Trees*. Available at `http://www.fi.muni.cz/~hlineny/Vyuka/GT/Grafy-lect-en-4.pdf` (2017/3/4).

[6] Jeremy Orloff and Jonathan Bloom. *18.05 Problem Set 2, Spring 2014*. Available at `https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/assignments/MIT18_05S14_ps2.pdf` (Spring 2014).

[7] Jeremy Orloff and Jonathan Bloom. *18.05 Problem Set 2, Spring 2014 Solutions*. Available at https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/assignments/MIT18_05S14_ps2_solutions.pdf (Spring 2014).

[8] Jeremy Orloff and Jonathan Bloom. *Conditional Probability, Independence and Bayes Theorem Class 3, 18.05, Spring 2014*. Available at https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/readings/MIT18_05S14_Reading3.pdf (Spring 2014).

[9] Jeremy Orloff and Jonathan Bloom. *Discrete Random Variables Class 4, 18.05, Spring 2014 Jeremy Orloff and Jonathan Bloom*. Available at https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/readings/MIT18_05S14_Reading4a.pdf (Spring 2014).

[10] Jeremy Orloff and Jonathan Bloom. *Discrete Random Variables: Expected Value Class 4, 18.05, Spring 2014 Jeremy Orloff and Jonathan Bloom*. Available at https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/readings/MIT18_05S14_Reading4b.pdf (Spring 2014).