MIT Introduction to Statistics 18.05 Slides 4 - Questions

John Hancock

February 22, 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.

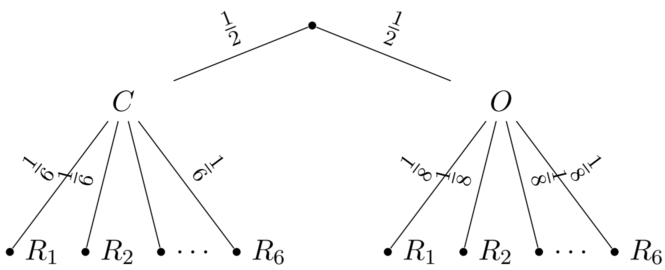
We are answering the questions that Orloff and Bloom ask in [3].
We use documentation in [5] to write LaTeXsource code for this document.

2 Conditional Probability of Unknown Die

The first question Orloff and Bloom give in [3] is:

> 1. The Randomizer holds the 6-sided die in one fist and the 8-sided die in the other.
> 2. The Roller selects one of the Randomizers fists and covertly takes the die.
> 3. The Roller rolls the die in secret and reports the result to the table.
>
> Given the reported number, what is the probability that the 6-sided die was chosen?

Note: we needed to see the solution in [4] in order to write the answer to this question.
We have two cases.
The first case is the Roller reports a 7 or an 8. Then the probability that the 6-sided die was chosen is 0.
The second case is the Roller reports a number with a value from 1 to six.
We draw a probability tree to get started on a solutionn. We refer to the section titled, "Shorthand vs. precise trees," in [1] for guidance on drawing the tree below.
C is the event that the Roller selects the cube shaped 6-sided die.
O is the event that the Roller selects the octohedron shaped 8-sided die.
$R1$, $R2$, ..., $R6$ are the events that the Roller reports one, two, or so on to six.



Since the probabilities on all edges in the tree connected to C are $\frac{1}{6}$, and the probabilities on all edges in the tree connected to O are $\frac{1}{8}$, we can calculate $P(C \mid R_1)$, and the result will be the same for any of the other leaf nodes in the tree above. This is because the calculation will involve the same numbers $\frac{1}{2}$, $\frac{1}{6}$, and $\frac{1}{8}$, and the same operations on these numbers.
Using the tree above, we can calculate $P(R_1 \mid C)$.
Now, we use Bayes' theorem in [1] to calculate $P(C \mid R_1)$

$$P(C \mid R_1) = \frac{P(R_1 \mid C)\, P(C)}{P(R_1)} \tag{1}$$

Now, we apply definitions for values on various parts of probability trees using the section titled "Shorthand vs. precise trees," in [1] to obtain values for the numerator and denominator on the righthand side of 1.
From the probability tree,

$$P(R_1 \mid C) = \frac{1}{6} \tag{2}$$

$P(C) = \frac{1}{2}$. Note: we are assuming the Roller uses either die with equal probability.
We apply Bayes rule [1] and the Law of Total Probability [1] to compute $P(R_1)$.

$$P(R_1) = P(R_1 \cap C) + P(R_1 \cap O) = P(R_1 \mid C)\, P(C) + P(R_1 \mid O)\, P(O) = \left(\frac{1}{6}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{8}\right)\left(\frac{1}{2}\right) \tag{3}$$

Now we have values for the numerator and denominator of the right hand side of 1.

$$P(C \mid R_1) = \frac{\left(\frac{1}{6}\right)\left(\frac{1}{2}\right)}{\left(\frac{1}{6}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{8}\right)\left(\frac{1}{2}\right)} \tag{4}$$

3 Concept Questions for CDF and PMF

In [3] Orloff and Bloom give us a table for a random variable X:

values of X	1	3	5	7
cdf $F(a)$	0.5	0.75	0.9	1

3.1 $P(X \leq 3)$

$P(X \leq 3) = 0.75$. We know this because of the definition of cdf, and because Orloff and Bloom give us the value of the cdf of X for $X \leq 3$.

3.2 $P(X = 3)$

We know

$$F(X \leq 3) = P(X = 3) + P(X = 1) \tag{5}$$

Therefore

$$P(X = 3) = F(X \leq 3) - P(X = 1) = 0.75 - 0.5 = 0.25 \tag{6}$$

4 Sum of Binomial Random Variables

4.1 Sum of Binomial Random Variables with Same Heads Probability

In [3] Orloff an Bloom pose the question that if $X \sim \text{binomial}(n, p)$ and $Y \sim \text{binomial}(m, p)$, what distribution does $X + Y$ follow?
In [2] Orloff and Bloom give the definition of a random variable that folllows a binomial distribution.
Therefore we apply the definition of a binomial distribution to the binomial random variables Orloff and Bloom give us in this section:
X is the number of heads in n independent Bernoulli trials, and Y is the number of heads in m independent Bernoulli trials.
For both X and Y we are given that the Bernoulli trials have the same probablity for success, p.
Therefore $X + Y$ is the number of successes in $n + m$ independent Bernoulli trials with a probability of success p.
Therefore

$$(X + Y) \sim \text{binomial}(n + m, p) \tag{7}$$

4.2 Sum of Binomial Random Variables with Different Heads Probability

We arrive at the answer by process of elimination.
In [3] Orloff and Bloom give us two random variables: X, and Z, where $X \sim \text{binomial}(n, p)$, and $Z \sim \text{binomial}(n, q)$.
Orloff and Bloom then ask us which distribution $X + Z$ follows, and they give us four options to choose from.
The first option is $\text{binomial}(n, p + q)$. This cannot be correct because $X + Z$ is a sum of the number of successes in $n + n$ independent Bernoulli trials.
The second option is $textbinomial(n, pq)$. This also cannot be correct because $X + Z$ is a sum of the number of successes in $n + n$ independent Bernoulli trials.
The third option is $\text{binomial}(2n, p + q)$. This cannot be correct because there is a counterexample.
We construct the counterexample: suppose $X \sim \text{binomial}\left(n, \frac{2}{3}\right)$, and $Z \sim \text{binomial}\left(n, \frac{2}{3}\right)$.
Then, if the third option were correct, $X + Z \sim \text{binomial}\left(2n, \frac{4}{3}\right)$. No probability can be greater than 1, so the third option cannot be correct.
This leaves us with the final option of, "other."

5 Number of Successes Before Second Failure

In [3] Orloff and Bloom ask us to describe the pmf of a random variable X where X is the number of successes before the second failure of a sequence of independent Bernoulli trials.
Let ω be a sequence of trials that fits the description of a sequence of trials that Orloff and Bloom give in this question.
Let Ω be the set of all ω.
We assume ω has $n+2$ trials, where n of the trials are successful, and two of the trials are failures.
Orloff and Bloom implicitly state that all the sequences of trials end in a failure because they are asking for the number of successes before the second failure.
Therefore the $(n+2)^{nd}$ element of ω is the second failure.
We can partition Ω into $n+1$ disjoint subsets containing one element each, where each subset has the first failure in a different position.
Therefore we can apply the Law of Total Probability [1] to comupte the probability of the first failure occuring in any of the $n+1$ positions in ω to be $(n+1)(1-p)$, where p is the probability of a successful Bernoulli trial in ω.
There are n independent successful trials in ω, with probability p, one unsuccessful independent trial in ω with probability $(n+1)(1-p)$, and one final unsuccesful independent trial in ω with probability $(1-p)$.
We know from [1] that the probability of the union of these independent events is equal to the product of the probabilities of the events. Therefore

$$p(\omega) = p^n (n+1)(1-p)^2 \tag{8}$$

6 Forgetful Geometric Random Variables

In [3] Orloff and Bloom as us to show that for a random variable X that follows a geometric distribution

$$P(X = n+k \mid X \geq n) = P(X = k) \tag{9}$$

Proof. We apply Bayes' theorem [1] to get started:

$$P(X = n+k \mid X \geq n) = fracP(X \geq n \mid X = n+k)\, P(X = n+k) P(X \geq n) \tag{10}$$

$P(X \geq n \mid X = n+k) = 1$. If we are given that $X = n+k$, then we arecertain that $X \geq n$.
Therefore we can rewrite equation 11:

$$P(X = n+k \mid X \geq n) = fracP(X = n+k) P(X \geq n) \tag{11}$$

□

References

[1] 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).

[2] 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).

[3] Jeremy Orloff and Jonathan Bloom. *Discrete Random Variables; Expectation 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/class-slides/MIT18_05S14_class4slides.pdf` (Spring 2014).

[4] Jeremy Orloff and Jonathan Bloom. *Discrete Random Variables; Expectation 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/class-slides/MIT18_05S14_class4_slides.pdf` (Spring 2014).

[5] ShareLaTex. *Typesetting quotations*. Available at `https://www.sharelatex.com/learn/Typesetting_quotations` (Spring 2014).