ECS 132 Fall 2020: Midterm Exam 1

November 10, 2020

Instructions

1. You may in no circumstances upload this midterm to private tutoring websites such as Course-Hero or Chegg. Remember all material related to this course is a property of the University of California and posting them is a violation of the copyright laws.
2. You may not upload the questions to any website and ask for help with the questions, nor may you send the exam to anyone else to solicit help.
3. The exam IS open-book, which means you may look up equations in the textbook, homework, discussions, course jupyter notebooks, or course notes
4. You may NOT help your peers or classmates under any circumstances. All questions should be directed to your TAs and professor
5. If you refer to a source (either a book or course notes), you must cite it.
6. You are MUST work these problems on your own - YOU MAY NOT WORK WITH PEERS.
7. Write your answers in GradeScope using markdown.
8. If you have questions, email your TAs and professor: sgburman@ucdavis.edu, dghosal@ucdavis.edu, davhaddad@ucdavis.edu, aacharyya@ucdavis.edu – You must email ALL of us, as only some of us will be responding at any given time. The subject of your email must begin with "ECS 132 Midterm Q: "

1 Problem

The background screen of your smart phone has 50 images, of which 10 are from your favorite destination, Putah Creek (PC). At the beginning of each day, your smart phone uses a shuffle feature to randomly order the images and display them over the course of the day. Note that after a shuffle, all 50 images are shown before they are re-shuffled the following day.

1. What is the probability that at least one of the first 5 pictures shown is a PC picture ?

Answer

We can follow the same argument as above. Let p denote the probability that at least one of the first 5 pictures shown is a PC picture. In order to find this we can find the complementary probability that none of first 5 pictures shown is a PC picture. We denote this probability by q which is $1 - p$. We can write down

$$\begin{aligned} q &= P(\text{none of first 5 pictures shown is a PC picture}) \\ &= \frac{{}^{40}P_5}{{}^{50}P_5} \\ &= 0.3105628 \end{aligned}$$

Thus

$$\begin{aligned} p &= 1 - q \\ &= 1 - 0.3105628 \\ &= 0.6894372 \end{aligned}$$

2 Problem

In Problem 1 you found the probability that at least 1 of the first 5 pictures shown is a PC picture on any given day. Using this probability, what is the probability that it will take more than 2 days for a PC picture to be among the first 5 pictures displayed?

2.1 Answer

On any given day, let p denote the probability that one of the first 5 pictures shown is a PC picture. This denoted as success. Then $q = 1 - p$ denotes the probability that none of the first 5 pictures shown is a PC picture. This is failure.

Let Y denote the number of failures until success. $Y \sim \text{Geo}(p)$ and the pmf is given by

$$P(Y = k) = (1-p)^k p \qquad k \in 0, 1, 2, 3, ...\infty \tag{1}$$

We need to find $P(Y \geq 2)$ and this is given by

$$\begin{aligned} P(Y \geq 2) &= 1 - P(Y = 0) - P(Y = 1) && (2) \\ &= 1 - p - (1-p)p && (3) \\ &= 1 - 0.6894372 + (0.3105628)(0.6894372) && (4) \\ &= 0.09644925 && (5) \end{aligned}$$

3 Problem

Suppose a pot contains 100 unique coins, of which 90 are *genuine* and 10 are *counterfeit*. We draw 5 of the 100 coins at random.

1. Compute the probability that all 5 of the drawn coins are genuine.
2. Write a simulation of this experiment using R, which calculates the probability. Your R script should repeat the experiment 1000 times, remember that each experiment is independent (you do not need to run your script).

Answer

1. Compute the probability that all 5 of the drawn coins are genuine.

   Let p denote the probability that all 5 of the drawn coins are genuine. Then p is given by

$$p = \frac{{}^{90}P_5}{{}^{100}P_5} \tag{6}$$

$$= 0.5837524 \tag{7}$$

2. Write a simulation of this experiment using R, which calculates the probability. Your R script should repeat the experiment 1000 times, remember that each experiment is independent (you do not need to run your script).

```
# strategy: We will consider numbers 1 - 90 as good and 91 - 100 are defect
# set.seed(1237)
m = 1000                    # number of samples to simulate
good = numeric(m)             # initialize for use in loop
for (i in 1:m)
 {
  pick = sample(1:100, 5, replace = FALSE)   # vector of 5 items from ith e
  good[i] = sum(pick <= 90) # number of genuine in ith experiment
 }
mean(good == 5)
```

4 Problem

Imagine that you are an Intrepid Insect Inspector. Last week, you identified a new species of beetle, which you named New Beetles. New Beetles look very similar to the previously described Blue Bitey Beetle. Today, while intrepidly inspecting, you find a community of beetles. You know with absolute certainty from past laboratory experiments on similar beetle communities that 62% of the beetles are New Beetles, and the rest are Blue Bitey Beetles. You also know that when doing identifications in the field, you misidentify Blue Bitey Beetles as New Beetles 52% of the time and you misidentify New Beetles as Blue Bitey Beetles 63% of the time.

1. What is the probability that you properly identify any beetle that you find in the field?
2. What is the probability that if you misidentify a beetle in the field, it is a Blue Bitey Beetle?

Answer

Consider the following events

N : A beetle is New Beetle
B : A beetle is Blue Bitey Beetle
I : Correct identification
M : Misidentification

Note that $P(B) = 1 - P(N)$ and $P(M) = 1 - P(I)$. We are given $P(N) = 0.62$, $P(M|B) = 0.52$, $P(M|N) = 0.63$.

1.

$$\begin{aligned} P(I) &= 1 - P(M) \\ &= 1 - (P(M|B)P(B) + P(M|N)P(N)) \\ &= 1 - (0.52)(1 - 0.62) - (0.63)(0.62) \\ &= 0.4118 \end{aligned}$$

2.

$$\begin{aligned} P(B|M) &= \frac{P(M|B)P(B)}{P(M)} \\ &= \frac{P(M|B)P(B)}{1 - P(I)} \\ &= \frac{(0.52)(1 - 0.62)}{1 - 0.4118} \\ &= 0.3359402 \end{aligned}$$