LAB: INTRO TO STAT ANALYSIS

Zeyi Qian

zeqian@clarku.edu

Office Hours: JC 201, Tuesday 4-5 PM & Thursday 3-4 PM

October 10, 2024

Question 1

A 2015 study conducted by the National Center for Health Statistics found that 51% of U.S. households had no landline service. We are going to pick five U.S. households at random:

- a) What is the probability that **all five** of them have a landline?
  $P(a\ household\ does\ have\ alandline) = 1 - 51\% = 49\%$
  $P(all\ five\ have\ a\ landline) = 0.49^5 = 0.028$
- b) What is the probability that **at least one of them** does not have a landline?
  $P(at\ least\ one\ of\ them\ does\ not\ have\ a\ landline) =$
  $1 - P(all\ five\ have\ a\ landline) = 1 - 0.028 = 0.972$
- c) What is the probability that **at least one of them does have** a landline?
  $P(at\ least\ one\ of\ them\ does\ have\ a\ landline) =$
  $1 - P(all\ five\ do\ not\ have\ a\ landline) = 1 - 0.51^5 = 0.962$

Question 2: a

The Mars company says that before the introduction of purple, yellow candies made up 20% of their plain M&Ms, red another 20%, and orange, blue and green each made up 10%. The rest were brown.

- a) If you pick an M&M at random, what is the probability that:
  - a. It is brown
    $1 - 20\% - 20\% - 30\% = 30\%$
  - b. It is yellow or orange
    $20\% + 10\% = 30\%$
  - c. It is not green $1 - 10\% = 90\%$

Question 2: b

- b) If you pick three M&M's in a row, what is the probability that:
  - a. They are all brown
    $0.3^3 = 0.027$
  - b. None are yellow
    $P(Not\ Yellow\ in\ one\ pick) = 1 - P(Yellow) = 1 - 0.20 = 0.80$
    $P(Not\ Yellow\ in\ three\ pick) = 0.80^3 = 0.512$
  - c. At least one is green
    $P(None\ Green) = 0.90^3 = 0.729$ $P(At\ least\ one\ Green) = 1 - 0.729 = 0.271$

Question 3

A survey asked students about their birth order and which college of the university they were enrolled in:

	birth order		
	1 or only	2 or more	Total
Arts and Sciences	34	23	57
Agriculture	52	41	93
Human Ecology	15	28	43
Other	12	18	30
Total	113	110	223

Question 3: a b c d

Suppose we select a student at random from this class. What is the probability that the person is:

- a. A Human Ecology student? $P(Human\ Ecology) = 43/223 = 0.19$
- b. A firstborn student? $P(Firstborn) = 113/223 = 0.5$
- c. Firstborn and a Human ecology student? $P(Firstborn\ and\ Human\ Ecology) = 15/223 = 0.6$
- d. Firstborn or a Human ecology student? $P(Firstborn\ or\ Human\ Ecology) = (113 + 28)/223 = 0.63$

Question 3: e & f

- e. A firstborn if you know they are a Human Ecology student?
  $P(Firstborn|Human\ Ecology) = P(Firstborn\ and\ Human\ Ecology)/P(Human\ Ecology) = (15/223)/(43/223) = 0.067/0.193 = 0.34$
- f. A Human Ecology student if you know they are a firstborn?
  $P(Human\ Ecology|Firstborn) = P(Firstborn\ and\ Human\ Ecology)/P(Firstborn) = (15/223)/(113/223) = 0.067/0.507 = 0.132$

Question 4

Suppose that 25% of people have a dog, 29% of people have a cat and 12% of people own both.

- a. What is the probability that someone owns a dog and not a cat?
  $P(Dog\ and\ No\ Cat) = P(Dog) - P(Dog\ and\ Cat) = 0.25 - 0.12 = 0.13$
- b. What is the probability that someone owns a cat or a dog?
  $P(Dog\ or\ Cat) = P(Dog) + P(Cat) - P(Dog\ and\ Cat) = 0.25 + 0.29 - 0.12 = 0.42$
- c. What is the probability that someone doesn't own any pets?
  $1 - P(Dog\ or\ Cat) = 1 - 0.42 = 0.58$
- d. What is the probability that someone we know owns a dog, also owns a cat?
  $P(Cat|Dog) = P(Dog\ and\ Cat)/P(Dog) = 0.12/0.25 = 0.48$

Question 64 from textbook

64. Shirts The soccer team's shirts have arrived in a big box, and people just start grabbing them, looking for the right size. The box contains 4 medium, 10 large, and 6 extra-large shirts. You want a medium for you and one for your sister. Find the probability of each event described.

a) The first two you grab are the wrong sizes.
b) The first medium shirt you find is the third one you check.
c) The first four shirts you pick are all extra-large.
d) At least one of the first four shirts you check is a medium.

Question 64 from textbook

- a. $P(\textit{first two wrong sizes}: \textit{not } M) = 16/20 \times 15/19 = 0.63$
- b. $P(\textit{third is } M) = 16/20 \times 15/19 \times 4/18 = 0.14$
- c. $P(\textit{all four } XL) = 6/20 \times 5/19 \times 4/18 \times 3/17 = 0.0031$
- d. $P(\textit{none } M) = 16/20 \times 15/19 \times 14/18 \times 13/17 = 0.38$
  $P(\textit{at least one } M) = 1 - 0.38 = 0.62$

Question 82 from textbook

82. No-shows An airline offers discounted "advance-purchase" fares to customers who buy tickets more than 30 days before travel and charges "regular" fares for tickets purchased during those last 30 days. The company has noticed that 60% of its customers take advantage of the advance-purchase fares. The "no-show" rate among people who paid regular fares is 30%, but only 5% of customers with advance-purchase tickets are no-shows.

a) What percent of all ticket holders are no-shows?

b) What's the probability that a customer who didn't show had an advance-purchase ticket?

c) Is being a no-show independent of the type of ticket a passenger holds? Explain.

Question 82 from textbook

$P(Advance\ purchase) = 0.60$
$P(Regular) = 0.40$
$P(No\ show \mid Regular) = 0.30$
$P(No\ show \mid Advance\ purchase) = 0.05$

- a. $P(No\ show) = P(No\ show \mid Regular)P(Regular) + P(No\ show \mid Advance\ purchase)P(Advance\ purchase) = 0.12 + 0.03 = 0.15$
- b. $P(Advance\ purchase \mid No\ show) = \frac{P(No\ show \mid Advance\ purchase)P(Advance\ purchase)}{P(No\ show)} = 0.2$
- c. Since $P(No\ show \mid Regular) = 0.3$ and $P(No\ show \mid Advance\ purchase) = 0.05$ are both different from $P(No\ show) = 0.15$, this means that being a no-show is not independent of the type of ticket. The probability of being a no-show changes based on whether the ticket is regular or advance-purchase.