**FINAL PROJECT: BEETOX DATASET**

**By: Raghvandersinh Solanki**

**April 30, 2023**

**Predata:**

**Total Species Social Bees: 273**

**Total Species Solitary Bees: 451**

**Total Social and Solitary Combined: 724**

**Total Species Oligolectic Bees: 370**

**Total Species Polylactic Bees: 348**

**Total Species Bees Captured: 146918**

**Total Paralyzed: 426**

**Total Active: 560**

**Imidacloprid**

**Chapter 2:**

**Section 2.3: A Review of Set Notation**

1. The Beetox dataset contains bees that were paralyzed with the imidacloprid and those who were not. P =paralyzed, B =Social Bees, S = Solitary Bees, A =Active Set A= {BP,BP,BA,SP} SetB = {BP,BP.SA,SA}

**Section 2.4: A Probabilistic Model for an Experiments: The Discrete Case**

1. I wanted to know how many species bees were “paralyzed(P) and active(A)” and “paralyzed or active”. From the database we know that there are 724 different species of bees being tested, , 58.8% bees were paralysed and 77.3% were not paralyzed.
   1. **, 43.1% were paralyzed and Active**
   2. **, 90% were either paralyzed or active.**

**Section 2.5: Calculating the Probability of an Event: The Sample-point method.**

1. There is an issue with these section and it assumes that we are working with small datasets, but fret not with the help of excel I can find an answer. In this case I wanted to find the Event of solitary bees(S) that had a floral specialization of Oligolectic(O). To find the answer I will be using the COUNTIFS method in excel and dividing it with the Sample Space which is 724
   1. With the COUNTIFS I found 370 were S and O and divided that with 724 and got .511 or 51.1% of the Solitary bees had a floral specialization of Oligolectic

**Section 2.6: Tools for counting sample points.**

1. In this case I wanted to find the different possible ways of picking a pair of social bees and solit
   1. **.** Jonathan and Nathan can pick out species of bees in 123123.

**Section 2.7: Conditional probability**

1. I wanted to create a little scenario using the data from the Beetox database. A group of scientists experimented on the 200 species of Solitary bees and 50 species of social bees and changed their color to blue. There is 37.7 %(A) chance to pick social bee and 18.3 %(B) that they are experimented on and 62.3 %(C) chance to pick a solitary bee and 44.3 %(D) that they are experimented on. What is the probability that we pick a solitary bee, given that they are experimented on? Is it Independent?
   1. Yes, P(C|D) = P(D)

**Section 2.10: The Laws of Total Probability and Bayes Theorem**

1. In 2050 scientists have created a device that lets you know the bee’s degree of floral specialization is either Polylactic or Oligolectic, but there is an 80%(F)chance that it will get an error. There is a 48.4%(P) chance that the bee is Polylactic and 51.6% (O) chance that the bee is Oligolectic. What is the conditional probability that the device detects a Polylactic bees?
   1. **20% that the detector detects the Polylactic bee.**

**Chapter 3:**

**Section 3.2: Probability of Distribution: Discrete Random Variables**

1. Thomas and Lincoln captured 10 bees (5 social and 5 solitary) and put them in a box. Lincoln bet Thomas that if he pulls out 1 social bee, he gets 1 dollar, if he pulls 2 he get 2 dollars so on and so forth. What are the chances he gets all 5 in a row?
   1. = 0.4% chance that Thomas wins all 5 games in a row

**Section 3.4: Binomial Distribution:**

1. Timothy’s house has 8 beehives in the wall and one day she was swarmed (luckily, she doesn’t have allergies), resulting in her calling pest control. The exterminator randomly starts removing the beehives. What are the chances that the bee’s hives have at least one solitary bee colony in them, assuming that all beehives are not the same.
   1. n = 9, p = 0.623: P(y >= 1) = 1 – P(0), P(0) = , P(y >= 1) = 1 – 0.0003 = 0.9997, There is a 99.97 % chance that there is at least 1 solitary bee colony

**Section 3.5: Geometric Distribution**

1. I wanted to see the chance of getting a paralyzed bee after 50 bees were not paralyzed
   1. p =0.588, q =0.412: , 0.0000000002% chance to get a paralyzed bee after 50 non paralyzed bees.

**Section 3.7: Hypergeometric Distribution**

1. Leonardo is known to have a box containing 20 solitary bees and 10 social bees, he randomly picks up 10 bees without replacement, what is the probability he exactly 10 solitary bees are selected from the box?
   1. N = 20 + 10 = 30, n = 10, r = 20: , 0.61% chance to pick up exactly 10 solitary bees without replacement.

**Section 3.8: Poisson Distribution**

1. Suppose a beekeeper checks the condition of a queen in the beehive. It takes him on average an hour to check the conditions of 10 queen bees. During the hour what is the probability he checks for at least 2 queen bees?
   1. , 99.9% chance that the beekeeper checks at least 2 bees within an hour.

**Section 3.11: Tchebysheff’s Theorem**

1. Suppose that a queen bee lays with an average size of .30 inches, with a standard deviation of .01. Find the lower bound of 100 queen bees that lay eggs between .26 and .34 diameter.
   1. , 92 eggs are expected to have ranges between .26 to .34 diameter knowing that the eggs have an average of .30 inches.

**Chapter 4:**

**Section 4.2 Probability Distribution: Continuous Random Variable**

1. Roberto loves dissecting, one day his friend picked 5 dead bees that contains 4 males and 1 female. Roberto puts the corpses into the box and randomly picks a bee to dissect. Let Y be the number of trials on which he dissects the female bee. Find the probability function Y and its corresponding distribution function.

**Section 4.3: Expected Values for Continuous Random Variables,** we would need an continuous function which we don’t have

**Section 4.4: Uniform Distribution.**

1. After looking into the Beetox dataset the bees were collected between the intervals of {59,337}. We can find the probability of bees being collected below 200.
   1. , 58.5% chance to collect bees between the interval of {59, 200}.

**Chapter 5:**

**Section 5.2: Bivariate and Multivariate**

1. We can use it to find the joint probability on how the imidacloprid affected the Social vs Solitary. and
   1. Let’s assume that there are 1000 bees, 700 social and 300 solitary. 300 social were paralyzed by the toxin and 400 were not; 200 solitary bees were paralyzed and 100 were not.

Y1

|  |  |  |
| --- | --- | --- |
| **0.3** | **0.2** | **0.5** |
| **0.4** | **0.1** | **0.5** |
| **0.7** | **0.3** | **1** |

**Y2**

**Section 5.3: Marginal and Conditional probability distributions**

1. We can continue off the pervious question and find the marginal Probability. Lets find the Marginal probability of P(Y1) and P(Y2)

P(Y1)

|  |  |
| --- | --- |
| p(0) | p(1) |
| 0.7 | 0.3 |

P(Y2)

|  |  |
| --- | --- |
| p(0) | p(1) |
| 0.5 | 0.5 |

**Section 5.4 Independent Random Variables:**

1. **Again, we can continue off the previous section.**