Final Project: Piece 1 Formal Report

Linda Mei

Stockton University

CSCI-3327-001: Probability and Applied Statistics

Professor Hoy Byron

December 9, 2022

Table of Contents

Introduction

The dataset that I chose was the “MyAnimeList Top 2000 Anime” dataset that I found from Kaggle. I chose this dataset because the data is fairly recent and has a lot of data while not being super overwhelming with too many attributes to focus on. I was looking for an interesting dataset that could be easy to understand and work with. Furthermore, it also contains a lot of information that I could utilize when making problems that involve probability and statistics.

Chapter 2

2.3 (6)

From the top 25 anime on MyAnimeList, 15 animes are “Action” anime, 9 anime are “Comedy” anime, and 3 anime are Slice of Life. Find the number of anime that is

a. “Action” but not “Comedy” anime.
    - There are 6 “Action” anime that are not comedy animes.
b. “Action” and “Comedy” anime.
    - There are 9 “Action” and “Comedy” animes.

2.4 (12)

Of the top 10 anime, it can be either the “Shounen”, “None”, or “Shoujo” demographic.

a. List the sample space for the top 10 anime.
    - S = {Shounen, None, Shoujo}
b. Find the probability of the top 10 anime being a “Shounen” anime
    - P(Top 10 anime is “Shounen”) = $\frac{8}{10}$ = 0.8

2.5 (28)

Five of the best anime are listed for two movie viewings for a friend’s movie night. One and only one anime has no demographic. The two movies for movie night are picked at random.

a. List the possible outcomes for this situation.
    - Let SH = Shounen, [SH1, SH2, SH3, SH4, None]
    - S = {(SH1, SH2), (SH1, SH3), (SH1, SH4), (SH1, None), (SH2, SH3), (SH2, SH4), (SH2, None), (SH3, SH4), (SH3, None), (SH4, None)}
b. Assign reasonable probabilities to the sample points.
    - P = $\frac{1}{10}$ = 0.1
c. Find the probability that the movie with no demographic is chosen for movie night.

- $\text{P} = \frac{4}{10} = 0.4$

2.6 (36)

A Netflix movie release has six TV anime series that can be watched in any order. How many different ways can these TV anime series be watched?

- $6nPr6$ = 720

2.7 (71 and Definition 2.10)

Consider two events, A and B. Event A represents that the top five anime was created by "Wit Studio" and B represents that the top five anime was created by "Cloverworks." $P(A) = 0.4$, $P(B) = 0.1$, and $P(A \cap B) = 0.1$.

a. Are events A and B independent?
    - $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.1}{0.1} = 1$
      $P(A|B) = P(A)$ is false 1 does not equal 0.4.
    - $P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{0.1}{0.4} = \frac{1}{4} = 0.25$
      $P(B|A) = P(B)$ is false 0.25 does not equal 0.1.
    - $P(A \cap B) = P(A)P(B) = 0.4 \times 0.1 = 0.04$
      $P(A \cap B) = P(A)P(B)$ is false 0.1 does not equal 0.04
    - Not any one of the above hold true, therefore, events A and B are dependent.

2.8 (94)

Recall two events, A and B. from question 2.7. There are two studios that produced some of the top five anime. The probability of a top five anime animated by "Wit Studio" is 0.4. The probability of a top five anime animated by "Cloverworks" is 0.1. The probability of a top anime animated by both "Wit Studio" and "Cloverworks" is 0.1.

a. Find the probability that a top five anime is animated by either "Wit Studio" or "Cloverworks" or both studios.
    - $0.4 + 0.1 - 0.1 = 0.4$
b. Find the probability that a top five anime is not animated by these two studios.
    - $1 - 0.4 = 0.6$

2.10 (124)

Context: Spy x Family centers around three main characters: Loid Forger (spy), Yor Forger (assassin), and Anya Forger (esper, mindreader).

A sample of Spy x Family (second most popular anime) anime fans contains 61% fans of Yor and 39% fans of Loid. It is reported that 36% of Yor fans and 64% of Loid Forger fans also like Anya Forger. A random anime fan chosen at random from this sample likes Anya Forger. Find the conditional probability that this fan is a fan of Yor.

- $P(Y) = 0.61, P(L) = 0.36,\ A = Likes\ Anya$
- $P(A|Y) = 0.36, P(A|L) = 0.64$
- Given: Likes Anya, Find: Y
- $P(Y|A) = \frac{P(A|Y)P(Y)}{P(A|Y)P(Y)+P(A|L)P(L)} = \frac{0.36\times0.61}{0.36\times0.61+0.64\times0.36} = \frac{0.2196}{0.45} = 0.488$

Chapter 3

3.2 (5)

A nerd and a normie have a child and they ask their child to match four anime titles to their respective anime posters. If the child assigns the four anime titles to the four posters, find the probability distribution for Y, the number of correct matches.

- 4 titles to match to a poster = 4! = 24 (need to draw out the 24 possibilities and choose one of the combinations as the correct "solution" and then do the probabilities below)
- $P(0) = \frac{9}{24} = 0.375$

  $P(1) = \frac{8}{24} = 0.333$

  $P(2) = \frac{6}{24}$ =0.25

  $P(3)$ = N/A

  $P(4) = \frac{1}{24} = 0.04$

3.4 (44)

The "ufotable" animation team is successful with a probability of p. Assume that the animation team animates four times and the results are independent of one another. What is the probability that exactly four are successful if p = .9?

- p (success) = .9, q (fail) = 1 - .9 = .1
- n (times animating) = 4, y = 4 (successful)
- p(y) = $4nCr4 \times (0.9)^4 \times (.4)^0 = 1 \times 0.6561 \times 1 = 0.6561$

3.5 (67)

In the anime, Blue Period, 20.1% of applicants to the Tokyo University of Art have undergone art prep school. Applicants are interviewed sequentially and are selected at random from the pool. Find the probability that the first applicant with art prep school experience is found on the sixth interview.

- p (success) = 0.201, q (fail) = 0.799
- P(Y = 6) = $(0.799)^5 \times 0.201 = 0.06545$

3.7 (102)

A bowl contains eight paper slips with anime genres written on them, of which three are "Action", six are "Fantasy", and two are "Comedy." Two paper slips are to be drawn from the bowl, one at a time without replacement. What is the probability that the two paper slips will be "Fantasy"?

- Action = 3, Fantasy = 6, Comedy = 2
- N = 8, n = 2
- r = 6, y = 2
- N - r = 2
- n - y = 0
- $\frac{6nCr2 \times 2nCr0}{8nCr2} = \frac{15 \times 1}{28} = 0.5357$

3.8 (122)

Customers arrive at the anime and figurines store in Akihabara, Japan according to a Poisson distribution at an average of six per hour. During a given hour, what is the probability that exactly nine customers arrive?

- P(Y = 9) = $\frac{6^9}{9!} \times e^{-6} = 0.0688$

3.11 (167)

The number of customers per day at the same anime and figurines store, Y, has been observed for a while and is found to have a mean of 16 and a variance of 2. Using Tchebysheff's theorem, find a lower bound for $P(11 < Y < 21)$.

- $\mu = 16, \sigma = 2$
- $P(11 < Y < 21) = P(11 - 16 < Y - \mu < 21 - 16)$

- $= P(|Y - \mu| < 5) \geq 1 - \frac{2^2}{5^2} = 0.84$

Chapter 4

4.2 (2)

Spy x Family episode one contains six different video links, but only one is working. The videos are randomly selected and tried, one at a time, until the video loads (videos that do not work are discarded before another is tried). Let *Y* be the trial number in which the video works. Find the probability function for *Y*.

- P(Y = y)
- P(Y = 1) = $\frac{1}{6}$
    - P(1st video fails) = $1 - \frac{1}{6} = \frac{5}{6}$
- P(Y = 2) = $\frac{1}{5}$
    - P(2nd video fails) = $1 - \frac{1}{5} = \frac{4}{5}$
- P(Y = 3) = $\frac{1}{4}$
    - P(3rd video fails) = $1 - \frac{1}{4} = \frac{3}{4}$
- P(Y = 4) = $\frac{1}{3}$
    - P(4th video fails) = $1 - \frac{1}{3} = \frac{2}{3}$
- P(Y = 5) = $\frac{1}{2}$
    - P(5th video fails) = $1 - \frac{1}{2} = \frac{1}{2}$
- P(Y = 6) = $\frac{1}{1}$
    - P(6th video fails) = $1 - 1 = 0$
- P(Y = 1) = $\frac{1}{6}$
- P(Y = 2) = $\frac{1}{5} \times \frac{5}{6} = \frac{1}{6}$
- P(Y = 3) = $\frac{1}{4} \times \frac{5}{6} \times \frac{4}{5} = \frac{1}{6}$
- P(Y = 4) = $\frac{1}{3} \times \frac{5}{6} \times \frac{4}{5} \times \frac{3}{4} = \frac{1}{6}$
- P(Y = 5) = $\frac{1}{2} \times \frac{5}{6} \times \frac{4}{5} \times \frac{3}{4} \times \frac{2}{3} = \frac{1}{6}$
- P(Y = 6) = $\frac{1}{1} \times \frac{5}{6} \times \frac{4}{5} \times \frac{3}{4} \times \frac{2}{3} \times \frac{1}{2} = \frac{1}{6}$

4.3 (28)

The proportion of time per day that all checkout counters in an anime store are busy is a random variable *Y* with a density function.

a. Find the value of *c* that makes $f(y)$ a probability density function.

   $f(y) = cy^6,\ 0 \le y \le 1,$

   $0,\ elsewhere.$

   - $\int_0^1 cy^6 dy = c\frac{y^7}{7} = F(1) - F(0) = \frac{1}{7} - 0 = \frac{1}{7}c$
   - $f(y) = 1 = \frac{1}{7}c$
   - $c = 7$

b. Find $E(Y)$

   - $\int_0^1 y(y^6)dy = \int_0^1 y^7 dy = \frac{y^8}{8} = F(1) - F(0) = \frac{1}{8} - 0 = \frac{1}{8}$

4.4 (45)

Upon studying low pay for Japanese animator contracts, "Kyoto Animation" (animation studio) finds that the animator contracts have low pay that is uniformly distributed between 10 and 20, in units of thousands of dollars (annually). Find the probability that the low pay on the next animation contract is below $15,000.

- $f(y) = \frac{1}{20-10},\ 10 \le y \le 20,$

  $0,\ elsewhere.$
- P(Pay < $15,000) = P(Y < 15) =

  $\int_{10}^{15} \frac{1}{10}dy = \frac{1}{10}y = F(15) - F(10) = \frac{15}{10} - \frac{10}{10} = \frac{5}{10} = \frac{1}{2}$

Chapter 5

5.2 (9)

Y_1 is the proportion of the "Demon Slayer" Nezuko plushies that are sold at the beginning of the week, and Y_2, is the proportion of "Demon Slayer" Nezuko plushies that are stocked during the week. Let Y_1 and Y_2 have the joint probability density function be given by

$$f(y_1, y_2) = k(3 - y_2),\ 0 \le y_1 \le y_2 \le 1,$$
$$0,\ elsewhere$$

a. Find the value of k that makes this a probability density function.

- $\int_0^1 \left(\int_x^1 k(3 - y)dy \right) dx = k(3y - \frac{y^2}{2}) = F(1) - F(x) = k\left((3 - \frac{1}{2}) - (3x - \frac{x^2}{2} \right)$
- $\int_0^1 k(\frac{5}{2} - 3x + \frac{x^2}{2})dx = \frac{5}{2}x - \frac{3x^2}{2} + \frac{x^3}{6} = F(1) - F(0) = \frac{5}{2} - \frac{3}{2} + \frac{1}{6} = \frac{7}{6}$
- $k(\frac{7}{6}) = 1$
- $k = \frac{6}{7}$

5.3 (23)

Y_1 is the proportion of the "Himouto! Umaru-chan" Blu-Ray anime CDs are stocked at the beginning of the week, and Y_2, is the proportion of "Himouto! Umaru-chan" Blu-Ray anime CDs that are sold during the week, given by

$$f(y_1, y_2) = 5y_1,\ 0 \le y_2 \le y_1 \le 1,$$
$$0,\ elsewhere$$

a. Find the marginal density function for Y_2.

- Let $x = y_1\ and\ y = y_2$
- $f_1(x) = \int_0^x 5x\, dy = 5xy = F(x) - F(0) = 5x^2$
- $f_2(y) = \int_y^1 5x\, dx = \frac{5x^2}{2} = F(1) - F(y) = \frac{5}{2} - \frac{5y^2}{2}$

5.4 (53)

In question 5.4, we derived that $f(y_1, y_2)$ is a valid joint probability density function. Are Y_1 and Y_2 independent?

a. $f(y_1, y_2) = 9(1 - y_2),\ 0 \le y_1 \le y_2 \le 1,$
$0,\ elsewhere$

- Let $x = y_1\ and\ y = y_2$
- $f_1(x) = \int_x^1 9(1 - y)dy = \int_x^1 9 - 9y\, dy = 9y - \frac{9y^2}{2} = F(1) - F(x)$

$= 9 - \frac{9}{2} - (9x - \frac{9x^2}{2})$

- $f_2(y) = \int_0^y 9(1 - y)dx = \int_0^y 9 - 9y\, dx = 9x - 9xy = F(y) - F(0)$
  $= 9y - 9y^2$
- $f(x, y) = f_1(x)f_2(y) = independent\ if\ true$
- $9(1 - y) = 9 - \frac{9}{2} - (9x - \frac{9x^2}{2}) \times 9y - 9y^2$
- Y_1 and Y_2 are dependent because they're not equal to each other

References

Correia, B. B. D. (2022, May). MyAnimeList Top 2000 Anime, Version 1. Retrieved December 9, 2022 from https://www.kaggle.com/datasets/brunobacelardc/myanimelist-top-1000-anime.