# Probability and Statistics Data Report

Dataset used: [Anime Dataset 2023 (kaggle.com)](https://www.kaggle.com/datasets/dbdmobile/myanimelist-dataset)

Additional Resource: [Anime Statistics and Facts By Country, Rating and Market Size (enterpriseappstoday.com)](https://www.enterpriseappstoday.com/stats/anime-statistics.html)

## Chapter 1

Five anime shows have ratings of 8.75, 5.67, 7.78, 6.65, 9.01 respectively.

1. Find the mean.

= 7.572

1. Find the variance.

= 1.593

1. Find the standard deviation.

= 1.262

## Chapter 2

Four anime shows are selected to be broadcasted at a special event. At the event there are two monitors where each monitor will play two of the four shows randomly selected. What is the sample space of the outcome?

Let A = Monitor 1, B = Monitor 2

[AABB, ABAB, BABA, BBAA, ABBA, BAAB]

A studio has hired 10 new voice actors for work. Three shows the studio funds require a voice actor. How many ways can the voice actor role be filled?

= =10! / 3! = 720

A convention holds a raffle to win merchandise where you spin a wheel. In the listing is five shirts, two figures, six keychains and four stickers. Suppose you can get lucky and win one of everything, how many different merch is there to win if you win one of everything from the raffle?

5(2)(6)(4) = 240

What is the probability of an anime being a movie given that it is a comedy and the two events are independent?

Let A = Comedy, B = Movie

P(A) = .09, P(B) = .18

P(B|A) = .18 = 18%

In total around the world, all of those who watch anime are 54% male and 46% female. 20% of male watchers and 80% of female watchers favor romance anime. A random person is chosen who enjoys romance anime, what is the conditional probability that this person is male?

Let M = Male, F = Female, R = Romance

P(M|R) = P(R|M)P(M) / P(R|M)P(M) + P(R|F)P(F) = .20(.54) / .20(.54) + .8(.46) = 22.6%

## Chapter 3

A user watches 10 shows and 7 of them are highly rated. The user is likely to rate a show high 24% of the time. Let Y be the number of shows the user has highly rated.

1. What is the probability of Y?

= 0.2%

1. The expected?

E(Y) = 10(.24) = 2.4

1. The variance?

V(Y) = 10(.24)(.76) = 1.824

The same user with a rating of 24% watches shows until they find one they enjoy. Let Y be the number of shows watched. What is the probability the user’s first show enjoyed is on the 4th try?

= 10.5%

Refer to the previous question, what is the probability the user will have two shows rated high on his fifth show?

= 10.11%

Using the mean and standard deviation results from problem 1, find a lower bound of 20 shows that are expected to have a rating between 6 and 10.

20(3/4) = 15

## Chapter 4

The number of shows paired with the genre Fantasy follows a Poisson distribution. With eight shows total, one was found to be a non-Fantasy show.

Y = Order show was watched. Y has a uniform distribution over the interval (0,8)

1. Find the probability that it was the first show.

P(0 < Y < 1) = 1/8

1. Find the probability that it was the last show.

P(7 < Y < 8) = 1/8

1. Given that the first four shows were indeed Fantasy, find the probability of the imposter being found as the fifth show.

P(4 < Y < 5 | Y > 4) = P (4 < Y < 5) / P( Y > 4) = (1/8)(1/2) = ¼

The Action genre’s ratings can be modeled as having an exponential distribution with a mean of 2.4. What is the probability of a rating exceeding 3?

P(Y > 8) = = .2865