Video Game Sales Report

Reece Cohen

CSCI 3327, Probability and Statistics

December 11, 2024

As a warning I have noticed some of the sales data of what I have is relatively old and some of it could be incorrect. An example of incorrect data I have is that the last time the page was updated was 8 years ago and the most recently released game on the chart is Imagine: Makeup artist with its release date listed as 2020. This cannot be correct since the page would have to be updated at most 4 years ago and search results list the game releasing in 2009. This data however is what I could find that is interesting to me.

**Chapter 2**

**Section 2.3**

**2.1** Suppose a family contains two children of different ages, and we are interested in the gender

of these children. Let F denote that a child is female and M that the child is male and let a

pair such as F M denote that the older child is female, and the younger is male. There are four

points in the set S of possible observations:

Let A denote the subset of possibilities containing no males; B, the subset containing two

males; and C, the subset containing at least one male. List the elements of A, B, C, A ∩ B,

A ∪ B, A ∩ C, A ∪ C, B ∩ C, B ∪ C, and C ∩ B.

**Revised**: Suppose we look at two different games of different release dates, and we are interested in who published them. Let N denote a game published by Nintendo and A denote a game published by Activision. A pair such as N A denotes that the older game is from Nintendo and the newer game is from Activision. There are four points in the set S of possible observations:

Let X denote the subset of possibilities containing no Activision games; Y the subset containing two Activision games; and Z the subset containing at least one Activision game. List the elements of X, Y, Z,

X = {NN}

Y= {AA}

Z = {NA, AN, NN}

There isn’t much to say about this question as it doesn’t make use that much of the data at hand.

**Section 2.4**

**2.10** The proportions of blood phenotypes, A, B, AB, and O, in the population of all Caucasians in the United States are approximately.41,.10, .04, and.45, respectively. A single Caucasian is

chosen at random from the population.

1. List the sample space for this experiment
2. Make use of the information given above to assign probabilities to each of the simple events.
3. What is the probability that the person chosen has type A or type AB blood?

**Revised:** The proportion of games released by publishers, Nintendo (N), Microsoft Game studios (MGS), Take-Two Interactive (TTI), and Sony Computer Entertainment (SCE), in the population of games with sales greater than 100,00 copies is approximately .04, .01, .02, .04, respectively. A single game is chosen at random from the population

1. List the sample space for this experiment
2. Make use of the information given above to assign probabilities to each of the simple events
3. What is the probability that the game chosen at random has either been published by Nintendo or Take-Two Interactive

There is a 6% chance that a game chosen from this list is published either from Nintendo or Take-Two Interactive

This question is interesting considering Nintendo has a large library of games and even with Take-Two Interactives respectable library size only having a 6% chance to get a game from those two publishers puts into perspective on how many games there are that have done well and even then, 6% is a lot for having games that have done well. I got the data for this question by taking a count of how many games total in the list 16600 and counted the games that were published under Nintendo (703), Microsoft Game Studios (189), Take-Two Interactive (413), and Sony Computer Entertainment (683)

**Section 2.6**

**2.35** An airline has six flights from New York to California and seven flights from California to

Hawaii per day. If the flights are to be made on separate days, how many different flight

arrangements can the airline offer from New York to Hawaii?

**Revised**: If I choose to buy a game for Xbox one from a list of 214 and then the next day I buy a game for ps4 from a list of 337. How many ways can I buy two games.

337 \* 214 = 72,118

I got the data for this query by filtering the games out by platform they are categorized on. I don’t really feel much from this answer, it is a lot of ways to buy just 2 games but this info does not do that much.

**Section 2.7**

**2.71** If two events, A and B, are such that P(A) = .5, P(B) = .3 and find the following

**Revised:** Two events, A and B, are such that P(A) = .86 which is the probability of choosing a game that was released at least by 2000 , P(B) = .66, which is the probability of choosing a game that was released between the years (inclusive) 1990 – 2010. And P B) = .11, which is the probability of choosing a game that was released between 2000 - 1990. Find the following

1. P(A|B) =

*Sources*

[Video Game sales By Gregory Smith on Kraggle](https://www.kaggle.com/datasets/gregorut/videogamesales)