PaSciRo Report

Lisa Luff

Figure 1: Image: https://www.tes.com/lessons/vru8sCaM4RJjOQ/star-trek

Contents

Abstract

The program I created is designed to take input from a user, and to run a simulation of a number of objects based on three distinct classes with differing attributes, moving around a grid and interacting with each other. This report is discussing the use, and mechanisms of the program, and how changing parameters can vary the output.

The parameters used are the start numbers of each object type, the dimensions of the grid, and how many times the simulation should run.

The outcome of this is an insight into how changing these parameters can alter the final outcome. While also demonstrating how much of an affect randomisation has on the diversity of outcomes.

My recommendations encourage further exploration of the current parameters, as well as shifts in the class variable values of each object class, and amendments to the elements being randomised to test how much these would affect diversity in output.

Background

The program is a simulation of a section of the universe named PaSciRo. It is inhabited by three alien species; Humans, Vulcans and Klingons based on the comic, book, TV and movie series Star Trek. Each species has unique characteristics based on my personal opinion of their characteristics from watching the tv series exclusively. The sector of the universe is represented as a 2-D grid where each row, and column is one lightyear in distance. Each run of the simulation represents one year passing.

Each time the simulation runs, the alien objects move around the grid at varying speeds based on their species, and have species specific likelyhoods of direction change. If they end up in the same grid location the aliens will either have babies, fight or do nothing. The likelyhood of each interaction outcome is affected by class specific attributes, and the genders of the two objects. Each class also has different fight skills, affecting the probability of who will win the fight, and who will lose. All fights are fights to the death.

Objects also have individually randomised lifespans, and will die and be removed from the simulation once they have existed in the simulation for that number of years. If there are no aliens left, or the number of aliens equal or exceed the number of grid spaces, the simulation will stop as civilisation would collapse from under, or over population. To avoid reaching these extremes, as there are fewer aliens, the probability to reproduce increase, due to the large amount of space and resources, while as room and resources become scarce due to reaching population capacity, they will have an increased probability of fighting each other for their share in it.

The parameters I chose give the user a large variety of ways to make significant change to the simulation, and it's output using the least number of inputs. I wanted the program to be easy to use, while allowing for substantial re-useability in the following ways:

- By allowing changes to how many of each specific species type, the user can explore how much of an inpact the species specific traits might alter the simulation.
- Altering the grid size allows the user to explore how the spread of the aliens might affect the simulation.
- By altering the length of time the simulation will run for, the user can explore how time might even things out, or push them to the extreme. Testing how the amendments they have made affect the lifespan of the civilisation.

Methodology

There are two main areas of change; the alien start numbers, and the grid size. I wanted to explore each of these independently. For each area I will be exploring:

- Start numbers:
    - How the different species unique attributes affect the simulation outcome
    - I assessed this by testing how their start numbers affect:
        * Simulation run time
        * Final species numbers and proportions
        * Species baby and fight numbers and proportions
- Grid size:
    - How spread affects the simulation outcome
    - I assessed this by testing how changing the grid size affects:
        * Simulation run time
        * Current total at simulation end
        * Current species proportions at simulation end

I kept the following equal:

- To give them all an equal chance to play out, I will keep the simulation run length equal across all experiments at 1000 years.
- Due to having numerous areas where randomisation occurs, I ran the simulation 5 times for each variation, with random.seed() set to 1111, 2222, 3333, 4444 and 5555 in all programs where randomisation is used to create consistancy, then find the average for comparison. Seeds are not usually present used in the simulation.

The baseline simulation all other simulations were compared to is:

```
$ python assignment_main.py 10 10 10 20 20 1000
```

I began with increasing each species start number to 20, and then to 50:

```
# Increase to 20
$ python assignment_main.py 20 10 10 20 20 1000
$ python assignment_main.py 10 20 10 20 20 1000
$ python assignment_main.py 10 10 20 20 20 1000

# Increase to 50
$ python assignment_main.py 50 10 10 20 20 1000
$ python assignment_main.py 10 50 10 20 20 1000
$ python assignment_main.py 10 10 50 20 20 1000
```

I then decreased the grid dimensions to 10 by 10, and increased it to 50 by 50:

```
# Decrease to 10 by 10
$ python assignment_main.py 10 10 10 10 10 1000

# Increase to 50 by 50
$ python assignment_main.py 10 10 10 50 50 1000
```

While the program does output a plot, for the purposes of this, I will be looking at the output data exclusively.

Results

Species Start Parameters

Looking at the results from changing the start number of species parameters first, below is a table of the averages from the 5 samples:

	Baseline	Human 20	Vulcan 20	Klingon 20	Human 50	Vulcan 50	Klingon 50
Runs	208.4	211.8	282.2	279.4	100.8	117.6	314.4
Current total: Human	234.6	181.4	64	23	289.2	120.2	71.4
Current total: Vulcan	10.6	134.4	179.2	147.6	60.8	231	69
Current total: Klingon	9.4	8.8	11.8	15.8	55.7	55.8	103.6
% of total: Human	0.6296	0.4405	0.20425	0.1579	0.7129	0.29304	0.18574
% of total: Vulcan	0.07526	0.33179	0.5368	0.4396	0.15027	0.5695	0.18061
% of total: Klingon	0.095109	0.02167	0.06004	0.20251	0.1368	0.13741	0.4336
Total babies: Human	623.8	542	230	91.6	875.6	368	276.6
Total babies: Vulcan	28.8	355.8	526.2	403.8	135.8	598.2	328.2
Total babies: Klingon	24	65.2	46.8	81	111.8	138.2	804.8
% of babies: Human	0.7142	0.52562	0.2926	0.26124	0.7932	0.33579	0.229049
% of babies: Vulcan	0.08724	0.32153	0.6199	0.41770	0.11105	0.5413	0.19008
% of babies: Klingon	0.19985	0.152852	0.08705	0.32096	0.09574	0.12296	0.5809
Total fights: Human	764.6	706	294	129.6	1170	462.6	356
Total fights: Vulcan	45.6	444.4	679.4	520.6	160.2	813.8	457
Total fights: Klingon	52	136	88.6	151.8	169.6	224.4	1527
% of fights: Human	0.6799	0.5141	0.2786	0.22967	0.7675	0.31151	0.18472
% of fights: Vulcan	0.13780	0.30381	0.5650	0.39897	0.16663	0.5436	0.16185
% of fights: Klingon	0.24515	0.18205	0.15638	0.3714	0.1286	0.1449	0.6534

Below are boxplots of the data, where Base, is baseline, and H, V or K +20 or +50 are which parameter was changed and by how much. H being Human, V being Vulcan and K being Klingon. These demonstrate the pattern of differences in the medians, as well as the variation.
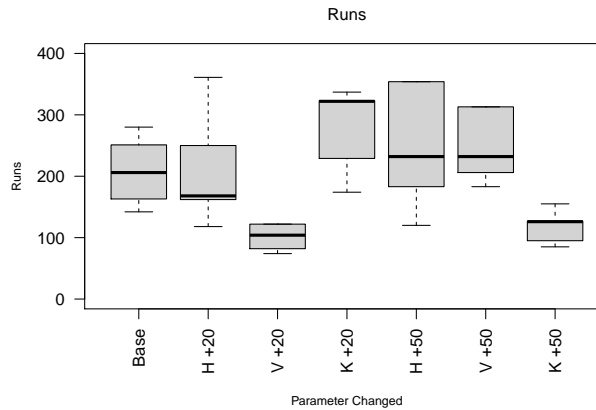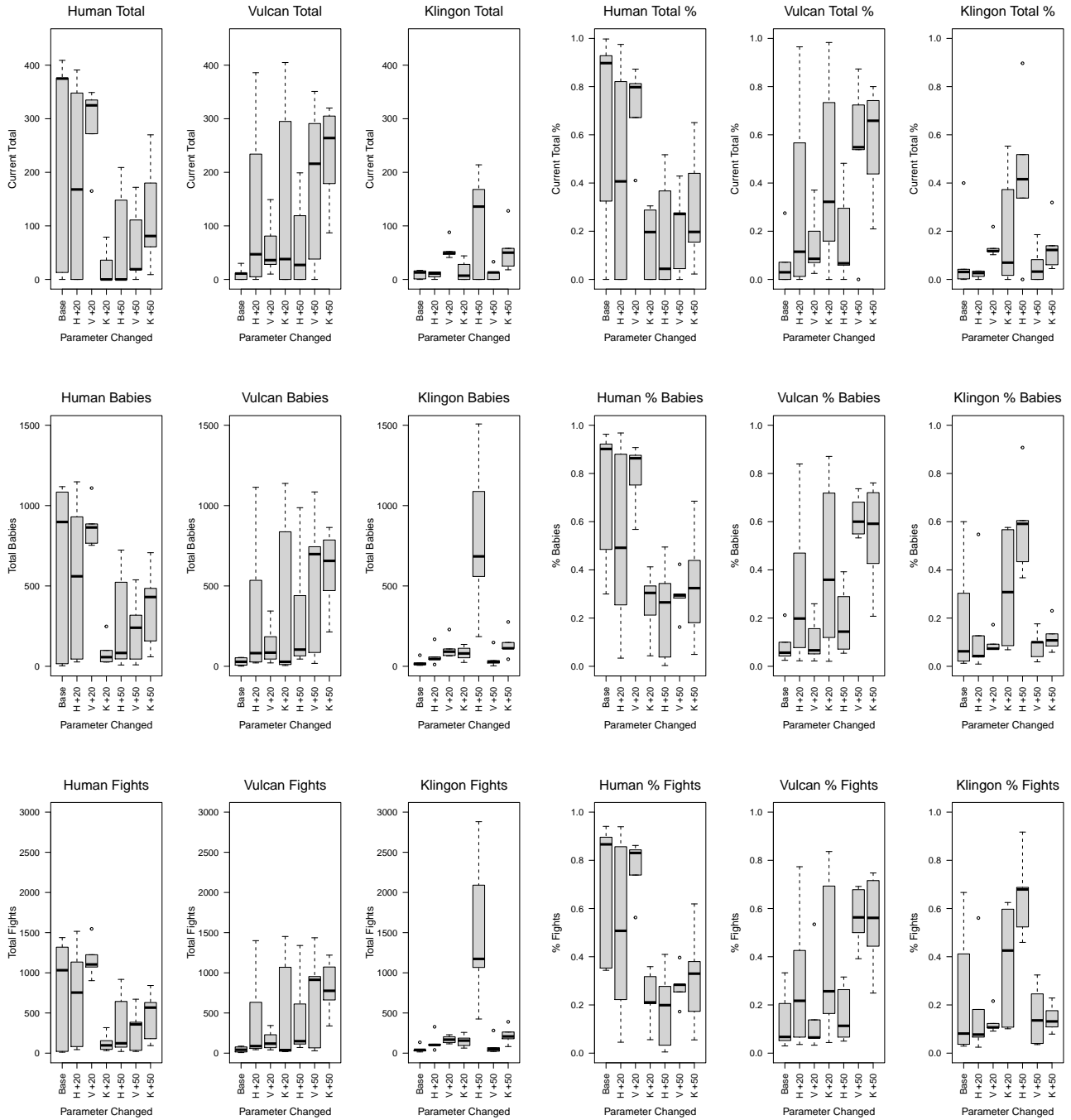
These boxplots clearly show that the data does not have equal variance, and is not Normally distributed. As the data sample is small, our hypothesis testing will need to be non-parametric.

I tested the null hypothesis that the means are equal regardless of the species start numbers parameter settings. Using the Kruskal-Wallis test, based on a significance level of 0.05, the outcomes are:

Measurement	P-Value	Significant
Runs	$P(\chi^2_{df=5} > 19.324) = 0.00365$	Yes
Human current total	$P(\chi^2_{df=5} > 11.155) = 0.0837$	No
Vulcan current total	$P(\chi^2_{df=5} > 9.387) = 0.153$	No
Klingon current total	$P(\chi^2_{df=5} > 14.158) = 0.02792$	Yes
Human % of total	$P(\chi^2_{df=5} > 10.712) = 0.0977$	No
Vulcan % of total	$P(\chi^2_{df=5} > 10.019) = 0.1239$	No
Klingon % of total	$P(\chi^2_{df=5} > 10.849) = 0.09314$	No
Human total babies	$P(\chi^2_{df=5} > 10.05) = 0.1226$	No
Vulcan total babies	$P(\chi^2_{df=5} > 9.4173) = 0.1514$	No
Klingon total babies	$P(\chi^2_{df=5} > 20.769) = 0.002019$	Yes
Human % of babies	$P(\chi^2_{df=5} > 14.174) = 0.02775$	Yes
Vulcan % of babies	$P(\chi^2_{df=5} > 15.131) = 0.01926$	Yes
Klingon % of babies	$P(\chi^2_{df=5} > 12.789) = 0.04652$	Yes
Human total fights	$P(\chi^2_{df=5} > 0.08818) = 0.08818$	No
Vulcan total fights	$P(\chi^2_{df=5} > 9.6457) = 0.1404$	No
Klingon total fights	$P(\chi^2_{df=5} > 20.785) = 0.022005$	Yes
Human % of fights	$P(\chi^2_{df=5} > 15.749) = 0.01517$	Yes
Vulcan % of fights	$P(\chi^2_{df=5} > 13.229) = 0.03853$	Yes
Klingon % of fights	$P(\chi^2_{df=5} > 12.8) = 0.04632$	Yes

The pure numbers look like there is significant change in the outcomes, however the hypothesis testing shows that more than half of the outcomes do not have an average outcome significantly different to the baseline tests average outcome.

The period of time the simulation was able to run for before collapse, was significantly affected by changes to start numbers of different species.

I hypothesise this is because the unique interaction probability attributes of some of the species in breeding and fighting would lead to a faster increase or decrease in population, when not challenged by the attributes of the other species.

The only species that was significantly affected by changes to the species number parameter was the Klingons, who saw a significant affect in all areas except for percent of total. The only time the Humans and Vulcans were affected was the percent of total babies, and percent of fights in Humans, which would make sense, as if one has a significant increase or decrease, it would affect the proportion of the total of all, if the others were not similarly affected.

I hypothesise that Klingons stand out due to their high fighting probability. As breeding requires an interaction between opposite genders, an equally high breeding probability, would not lead to an equal amount of babies and fighting. So a high fight probability would have more of an affect on interactions than a high breeding skill.

I hypothesis that the run number is significantly affected due to the attributes too, as a particularly large probability in breeding or fighting would have an exponential effect as populatino changes, and would lead to rapid extinction or civilisation collapse from over population.

Grid Size Parameters

Below is the table with the averages of the outputs from changing the grid parameters:

	Baseline	Grid -10	Grid +50
Runs	208.4	183.6	148.8
Current total	246.6	61.2	0
Human % of total	0.6296	0.2840	0
Vulcan % of total	0.7526	0.1264	0
Klingon % of total	0.095109	0.1896	0

Below are the boxplots to explore the pattern of difference in the medians, and the variance with changes in grid parameters. Again, Base represents the baseline data, while G -10 represents the 10 by 10 grid, and G +50 represents the 50 by 50 grid:
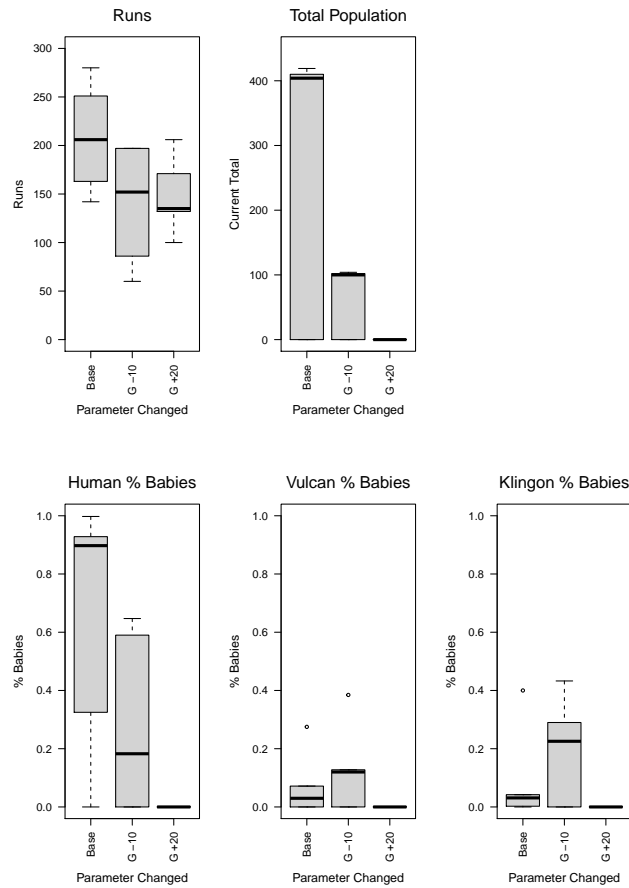
Once again the barplots show that the variances are not even, and that the data is not Normally distributed. So due to the small sample size the hypothesis testing needs to be non-parametric.

I tested the null hypothesis that the means are equal regardless of grid parameter changes. Using the Kruskal-Wallis test, based on a significance level of 0.05, the outcomes were:

Measurement	P-Value	Significant
Runs	$P(\chi^2_{df=2} > 2.3792) = 0.3043$	No
Current Total	$P(\chi^2_{df=2} > 4.8109) = 0.09022$	No
Human % of total	$P(\chi^2_{df=2} > 6.8176) = 0.03308$	Yes
Vulcan % of total	$P(\chi^2_{df=2} > 4.4545) = 0.1078$	No
Klingon % of total	$P(\chi^2_{df=2} > 5.4059) = 0.06701$	No

Looking at the table, you can see that the larger grid always runs out of alien inhabitants, however, only the Human percentage at end of the simulation has a significant difference.

I hypothesise that this is due to the smaller grid being so similar to the baseline grid, cancelling out to a reduced significance when comparing all three. As there is a definite trend evident in the raw means in the table, and the barplots, that as the spread of the aliens increase, the civilisation will be more likely to collapse due to the population dying out. I also hypothesise that as spread increases, the speed at which this happens will increase.

I hypothesis that the reason only the humans were significantly affected is because they have the highest breeding probability, and therefore usually have lots of babies. So when large spread means they're unable to find mates, there will be a larger drop simply due to there being a greater propensity to decrease.

Conclusion

My conclusion based to the number of factors significantly affected by the parameter changes, and the patterns of these, that what affects the simulation outcomes the most is the interaction probability attributes of the alien species.

Changing the proportions of starting aliens only leads to significant increases or decreases if their attributes are able to affect interaction probabilities strongly enough.

Changing the grid slightly does not have much affect, however if species numbers don't increase proportionately, the speed at which civilisation has no population left is faster as it increases.

Future Work

I believe that future work would benefit from exploring a greater range of spread using the grid size parameters to find at what point the simulation is significantly affected.

The real curiosity though would be exploring changing alien attributes. When I first created the program I had a non-binary gender, and instances who did not breed, however the civilisation was running low on population too fast, so I removed these. It would be interesting to see how the outcomes would change as attributes were increased and decreased, but also as attributes are added or removed.