y1 = np.exp(x)

# Proctor and Gamble Servo Motor Line Analysis

Ryan Day, Parker Hamilton, Cody Kesler, Spencer Yu

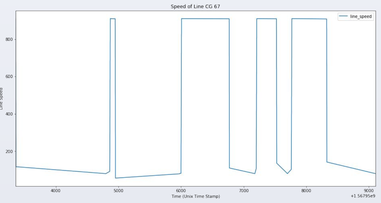
Theory of Predictive Modeling

## Problem Statement

Procter & Gamble have manufacturing plants worldwide with multiple diaper manufacturing lines in each of them. These lines usually run for multiple hours at a time, but if an anomaly or failure occurs that affects diaper output quality, the operators stop the line and fix the problem before restarting. This decreases output and overall line effectiveness, and so is important to examine. We received data from P&G related to the servo motors that operate these lines. We use this data to create predictive models that characterize and predict manufacturing line effectiveness.

## Background and data

Run time for a line is given in the line_speed feature of the data. The graph on the right shows 4 different run times of varying length. The ramp-up time is when the line speed is increasing, the run-time is then the plateau at 910 and the ramp down time is the time when the line is decreasing. Short run times are a problem for the factory as a loss of 20-100 diapers is taken for every ramp up and down. This becomes a large net loss for Proctor and Gamble as running many factories with multiple lines per factory adds up to many wasted diapers.




Line CG 67 Line Speed data

There was 6 data files provided, one for 6 different line in 3 different factories around the world. Each data file contained 5 measurements from the feeder motors of 10-11 different machines on each line. These measurements are Position Error, Actual Velocity, Commanded Velocity, Torque, and Position Error. The different machines are denoted as ALU, CUU, etc. There was also Line state, and Line Speed given in the data. The data was all time series data taken at a minute interval in a 24 hour period.

| | Tag Name | Value_Double | time |
|---|---|---|---|
| 6583756 | f_CUU.Meter.Omega.Axis.PositionError.Actual | -0.024774 | 1.568684e+09 |
| 3069370 | f_ALU.Meter.Omega.Axis.Velocity.Actual | 4940.984000 | 1.568679e+09 |
| 1640990 | f_ALU.Meter.Omega.Axis.Torque.Actual | 13.000000 | 1.568362e+09 |
| 6482663 | f_CUU.Meter.Omega.Axis.PositionError.Actual | 0.000000 | 1.568126e+09 |
| 6670218 | f_CUU.Meter.Omega.Axis.PositionError.Actual | 0.024023 | 1.569154e+09 |
| 11138718 | Line.State | 1.000000 | 1.568584e+09 |
| 10136760 | f_TU.Meter.Omega.Axis.Velocity.Actual | 7513.575000 | 1.568057e+09 |
| 1092438 | f_ALU.Meter.Omega.Axis.PositionError.Actual | -0.000751 | 1.569655e+09 |
| 9883174 | f_TU.Meter.Omega.Axis.Torque.Actual | 12.600000 | 1.568633e+09 |
| 1727912 | f_ALU.Meter.Omega.Axis.Torque.Actual | 9.700000 | 1.568530e+09 |

Sample of Line CG 67 data

## Main Objectives

1. Define a measure of effectiveness for each production line to compare line production based on the following features: Total up-time, number of ramp-up's and ramp-downs for each line, and the length of up-time for each line.
2. Predict if a run will last longer than a given tolerance level based on the first 10 minutes of run time data.

## Problem 1: Defining Line Effectivness

### Method

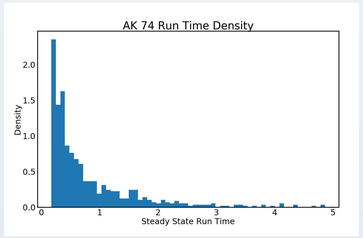
To complete the first objective, we used maximum likelihood estimation and statistical tests. There were six lines, two in each factory, that we examined and compared to one another. We parsed through the data in order to identify when a line ramped up, when it ramped down, and the duration of the run. We analyzed the distribution of durations of times that the run was active for. The distributions had many short run times and few long run times, and so were very skewed, and looked like they could be represented by an exponential distribution.

Because of this, we decided to characterize the distributions by an exponential distribution. The single exponential parameter tells us how much probability mass is located at longer run times between distributions. We used an MLE parameter estimate to describe the behaviour of each distribution and found that different plants had significantly different steady state run time distributions. The lines from the AK plant had the best distribution, with the longer run times making a larger portion of the probability mass, while the CG lines performed the worst.

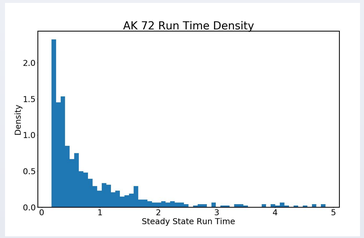
In order to tell if the distributions were significantly different than each other, or within the realm of statistical noise, we used a Mann-Whitney U test on each pair of lines (also known as the Wilcoxon rank-sum test). We chose this test because it does well on data that is not normally distributed. We used a Bonferroni Correction to adjust our significance level, since with so many statistical tests it makes it more likely to have a Type 1 statistical error (Falsely rejecting the null hypothesis)

### Results

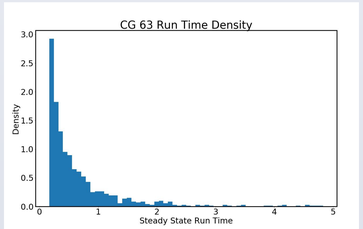
There were four pairings of the Mann-Whitney U test that were significant. These were between the AK lines and the CG lines. They all had p-values < 0.00333, which was our Bonferroni corrected significance level (0.05/15 statistical tests performed). From this it is strongly suggestive that there are differences between the AK lines and CG lines. We could not reject the null hypothesis from the other pairings (EUS and CG, EUS and AK) that they were significantly different from each other.



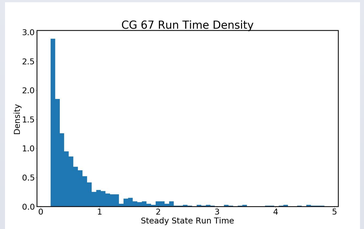

Plot of the density of steady state run times for the AK 74 line. The parameter fit for the exponential MLE was $0.66 \pm 0.05$




Plot of the density of steady state run times for the AK 72 line. The parameter fit for the exponential MLE was $0.62 \pm 0.04$




Plot of the density of steady state run times for the CG 63 line. The parameter fit for the exponential MLE was $0.50 \pm 0.03$




Plot of the density of steady state run times for the CG 67 line. The parameter fit for the exponential MLE was $0.51 \pm 0.03$

## Problem 2: Predicting Line Up-time

### Method

To accomplish our second objective, we developed a binary classifier neural network for predicting if a run time will be greater than a certain threshold. The input data is the the first 10 minutes of run time data. Each data point is a ramp up time for which the up-time was longer than 10 min. The output label is True/False for longer or shorter than the set tolerance level. The data includes temperature, torque, position error, and actual velocity - commanded velocity. This data is one minute frequency so there are 10 data points for each measurement with 11 machines per line. This gives 440 data points per entry and there are about 350-420 entries for each line. Since we have data for 6 lines there is roughly 2400 entries of length 440 to use to train the network.

### Results

At the beginning we did not expect very good performance from the neural net due to a lack of data. Not only is 2400 rows relatively small for a neural net, our data for each of the ramp ups was sampled at a minute frequency. Hence, if there was some indicator of failure, in order for it to be picked up by the neural net it would have to have a significant and reproducible impact on servo performance over a time period of 1 minute or greater. If that was not the case, then there is no way to predict failures.

We found our expectation to be the case as we could not get the classification accuracy above 50%. In order to combat the low number of rows we attempted to use various higher learning rates, but it ultimately had no effect on classification performance. We additionally found that due to positive/negative tag imbalance, our model tended to label all the test data as negative. We attempted to rectify this by re-balanced the training data set to contain equal amounts of failed and successful ramp ups. However, this only managed to produce a few positive ratings. By altering the training metric to only consider false positives and false negatives (and still using a balancing training set), we managed to produce more balanced ratings (i.e. not all negative labels). However it had a negative effect on overall accuracy. This is a worthwhile trade off due to the uselessness of a model that only gives one answer. Finally, we found that net architecture also had no effect on model performance.

An example classification matrix using a rebalanced training dataset and a false negative metric:

| | Predicted Failure | Predicted Success |
|---|---|---|
| Actual Failure | 116 | 252 |
| Actual Success | 28 | 56 |

Accuracy: .38 Sensitivity: 0.67 Specificity: .32

Positive Predicted Value: 0.81

Negative Predicted Value: 0.81

Due to the failure of the black box modeling attempts we believe our original hypothesis that the servo data is not rich enough to drive effective servo ramp-up classification.

## Future Work

### Objective 1

The method used to score lines is dependant on the assumption of an exponential distribution. A larger sample of run times from different plants would allow a better investigation of the validity of this assumption. A larger dataset from more plants at different times, perhaps along with product production data, would allow a more informative analysis of the MLE parameter fits. This would allow us to both find if theses parameter differences are significant as well as using product production data to determine if this scoring method is informative. Given data on servo failures we could also compare distributions correlating to servo failures.

### Objective 2

In terms of predicting up-time success, future work will need to include near-continuous time data as servo metrics collected at 1 minute collection intervals were too sparse to capture the indicators for failure and success. Second/half second intervals at all times for all metrics will allow for greater balance in the data for predicting up-time success as well as faster results. Knowing whether or not a line will run longer than an hour and a half after 10 minutes of up time is not an extremely useful prediction, however, if there were second time frequency data we would be able to predict successful up-times within a much smaller time window of ramp-up.

Predicting the length of the run time given the ramp up state, history of the line, and up-time periodicity, would also be of use in warning operators of potential quick failures of the line and the most unproductive times of the day for a given line. It would provide the ability to understand where to focus maintenance and effort on the lines.