Submitted March 6, 2015

**Title: Profiling Sex Differences in Openfield Behavior**

Capstone Project for Python Data Science Course

Thinkful <https://www.thinkful.com>

**Author:** Anna Delprato, Ph.D.

BioScience Project, Wakefield MA 01880 US

[delprato@bioscienceproject.org](mailto:delprato@bioscienceproject.org)

tel:(877) 339-5703

INCIA, Université de Bordeaux 1

avenue des Facultés, 33405 Talence FR

tel :+33 (0)5 40 00 89 00

**Hypothesis**

Identifying patterns in behavior may be useful for understanding how individuals experience the environment and process sensory input. In this study we attempt to profile, classify, and predict behavioral traits in mice that are associated with the openfield challenge. Our working hypothesis is that there are male/female differences for at least some of these traits.

**Background**

The openfield is a classic behavioral test that is used to measure anxiety, exploratory behavior and general locomotor activity in rodents. The animal is placed in a well lit, open environment which is intended to mimic a situation of predator vulnerability. The test is designed this way because rodents have an innate tendency to avoid bright lights and open spaces. Typically, at the start of the test, the animal remains at or travels along the periphery of the openfield apparatus. After several minutes, the apparent threat of danger is past and the animal becomes habituated to the environment and begins to explore.

**Methods**

**Openfield test**

The test subjects are monitored continuously throughout the experiment with video tracking software and also by an individual that manually registers the frequencies of certain features. Each experiment lasts 20 minutes. The parameters measured are: locomotion (total distance traveled), thigmotaxis (time spent at the periphery/thigmotaxis zone (surface within 5 cm of the walls), thigmodtaxis (distance from the center), grooming duration, and the frequencies of leaning (standing upright on hind legs with one or two forepaws against the wall), rearing (standing on hind legs with both forepaws free), jumping, grooming, and defecation.

**Data analyses**

The data set used in the analyses is derived from a broader quantitative genetics study that involved several behavioral experiments designed to identify and map heritable traits using the BXD recombinant inbred mouse strains. For the openfield challenge, 53 BXD recombinant inbred strains along with the two parental lines were tested. The values in the data set considered here represent the strain means. In most cases, the strain number was 10 for each sex. In this study, the data is used to profile sex differences only. Strain effects are not evaluated here.

For the analyses, the data were formatted as csv files and read into Python (https://www.python.org) as Pandas dataframes (http://pandas.pydata.org). The general statistical tests, mean, variance, range etc and Spearman Rank correlations were performed with Pandas. The histograms and scatter plots were created with Matplotlib (http://matplotlib.org). For the student’s t-test (http://www.scipy.org), Support Vector Machine (SVM) and Gaussian Naive Bayes (NB) analysis, the Pandas dataframes were converted to Numpy arrays (http://www.numpy.org ) prior to further processing. The SVM and NB analyses were done with Scikit learn (http://scikit-learn.org) using the default parameters (SVM: Linear kernel, C=1) for those algorithms. The SVM results were visualized with methods adapted from https://github.com/jakevdp/sklearn\_scipy2013.

.

**Results**

***Distributions and relation between variables by sex***

Histograms and scatter plots were used to evaluate the distribution properties and the relationship between variables for each sex. In general, the data are normally distributed and have relatively the same size tails with some variation. In some cases, there is positive or negative skewing but this is slight. Each of the scatter plots show a positive relationship between males and females for all of the features. The scatter plots shown in Figure 1, panels A, B, and C have significant clustering whereas, the data points in the scatter plot shown in Figure 1 D are disperse and do not exhibit a tight banding pattern.

**Statistics**

*Mean, SD, Variance, and Range*

The Mean, SD, Range, and Variance for both sexes are comparable or within 15% for locomotion, grooming frequency, thigmotaxis, and rearing. For leaning, male and female data have similar Means, SD, and Variance but the Min value is lower by ~18%. The statistics for grooming duration differ the most between the sexes. The Mean is ~20% higher and the SD ~38% greater for females. The Max value is also ~30% higher for females.

**Correlations**

Spearman rank correlation matrices were calculated to identify features within male/female data that covary with each other as well as among the male/female data to identify the lowest correlation values that may be useful for classification and prediction. The Spearman rank correlation method is preferred because it is not sensitive to the distribution Mean or to outliers.

Significant Spearman correlations within the male dataset are observed between the variables, grooming duration and locomotion which negatively covary (p=-0.659086). For females, the closely related variable, grooming bout, also negatively covaries with locomotion (p=-0.612463). This is a reasonable association because the more grooming that occurs, the less distance traveled.

Between the male and female datasets, the correlations are all positive and all significant (p ≥ 0.5) ranging from p=0.6-0.8. The lowest correlations observed are for rearing: p=0.609, grooming frequency: p=0.613, and defecation: p=0.600.

**Student’s t-test**

The Student’s t-test was used to determine whether there is a significant difference between males and females for the rearing, grooming frequency, and defecation features. The p-values obtained for each of these variables are smaller than the threshold values of 1%, 5% or 10%. The p-values for the rearing and defecation variables are lowest (t(1) l2.57l, p=0.01 and t(1) l1.90l), p=0.002) whereas, for grooming frequency, the resulting p-value is above the 5% threshold value (t(1) l3.22l, p=0.06).

**Machine learning**

The traits with significant or near significant p-values (rearing, grooming frequency, and defecation) identified from the Student’s t-test were evaluated with machine learning algorithms to see whether these traits could be used as predictors to discern male from female behavior in an openfield test.

**SVM**

The SVM classifier attempts to partition the data to maximize the distance between the data points and the margin boundary or the buffer between them. For these data, none of the possible combinations of the three traits partitioned cleanly with SVM. While many of the points are mis-classified, a clear separation for many of the points is also observed. For each set of features, the number of mis-classified points on either side of the boundary line is about equal (17/18) but is slightly skewed for the rearing and defecation combination which is around 20/16. The total of mis-classified points represents ~1/3 of the data.

**Naive Bayes**

The NB classifier takes an observation, calculates the Bayesian probability of each possible hypothesis, and chooses the hypothesis with the highest probability (the Maximum A Posteriori (MAP) hypothesis) to classify the observation.

For the rearing, grooming frequency, and defecation data, parameters could be defined where the NB prediction successfully discriminates between male and female. However, as seen with the SVM analysis ~ 1/3 of the data is mis-classified.

**Conclusions**

Overall, male/female performance in the openfield test is highly similar, however differences were observed for several of the features (rearing, grooming frequency and defecation). Based on preliminary results obtained with the machine-learning algorithms, NB and SVM, it may be possible to use these features in a classification/prediction scheme. Additional data and more homogeneous populations i.e. where strain effects are either considered or filtered out, may or may not help to reduce the number of mis-classifications. Biological data is fuzzy in general and this is especially true for behavioral data which has a great deal of variation. Fine tuning the parameters for the machine learning algorithms could produce a better classification and/or implementing other machine-learning algorithms may be more amenable to profiling this type of data.