COS 301 Team ANTZ DriveStats Algorithm

Written By Axel Ind 12063178

Background

The algorithm of the DVT Drivestats application returns a score from 0.0-10.0 as an evaluation of the driving ability of the driver in question.

This score is calculated based on sensor input regarding a particular trip the user took. The sensors involved are:

- GPS coordinates
- Speed
- Accelerometer (one for each of the x, y, z axes)

Implementation

Steps

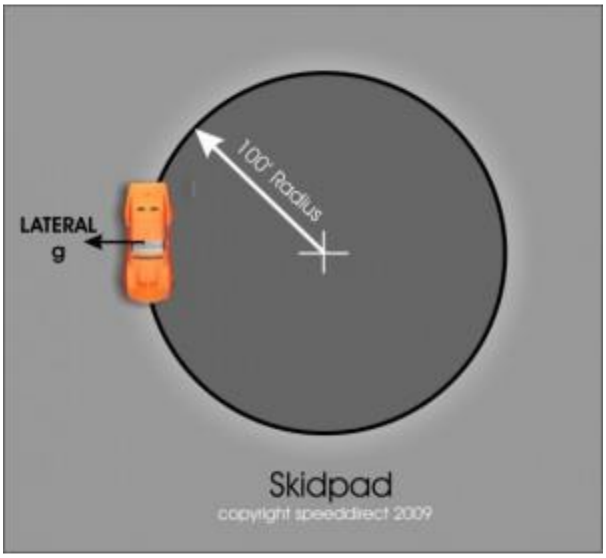
1. Determine how fast constitutes bad acceleration (4.2m/s/s forward, 30% lateral G – meaning 30% of normal earth gravity meaning about 3m/s/s, 2ms/s/s up or down) (Elert, 2001) (Anon., 2015).
2. Determine how many bad things all previous users have done per second.
3. Determine how many bad things a person does per second.
4. The data gathered – specifically the number of bad things per second of all drivers, and for our current driver - follows a layout called a Poisson distribution; because it expresses the probability of a given number of events occurring in a fixed interval of time and/or space if these events occur with a known average rate (Wikipedia, 2015).
5. Due to the nature of the distribution observed, we can approximate out Poisson distribution to a normal distribution. The normal distribution is useful because it has several well-described properties.

   > For sufficiently large values of λ, (say λ>1000), the normal distribution with mean λ and variance λ (standard deviation $\sqrt{\lambda}$), is an excellent approximation to the Poisson distribution. If λ is greater than about 10, then the normal distribution is a good approximation if an appropriate continuity correction is performed, i.e., P(X ≤ x), where (lower-case) x is a non-negative integer, is replaced by P(X ≤ x + 0.5).
   >
   > $$F_{\text{Poisson}}(x;\lambda) \approx F_{\text{normal}}(x;\mu=\lambda,\sigma^2=\lambda)$$

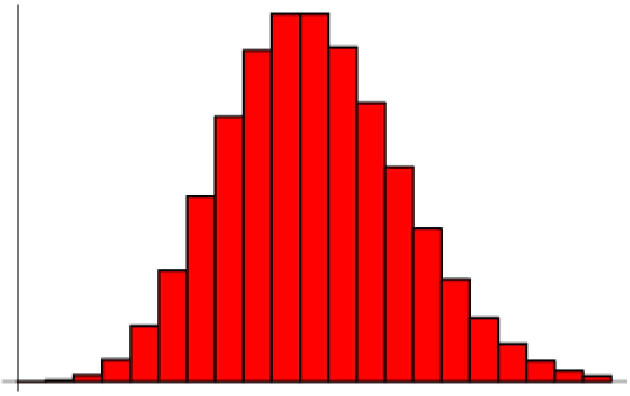
6. Using this information, we are able to estimate a probability that normal distribution of the current driver was a result of random chance when compared to the normal distribution of all drivers.
7. One of the useful properties of the normal distribution we exploit, is that we are able to estimate a position on a normal distribution for the likelihood of the observed results, and calculate its area.
8. Because the area under a normal distribution always equals 1, we can simply multiply the area by 10 to achieve a score out of 10 for the quality of that person’s driving.

Diagrams

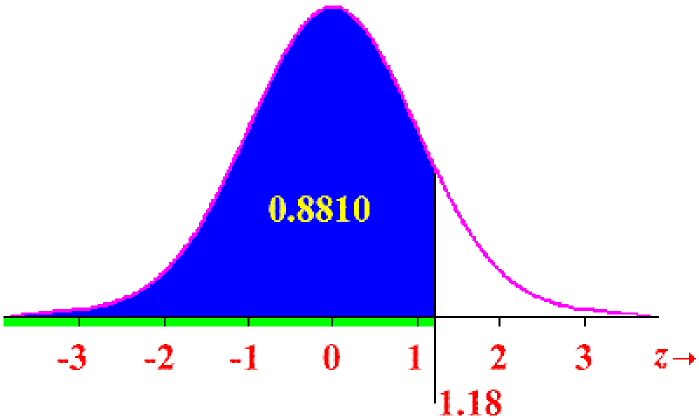
Cornering force



Typical Poisson distribution



Example area under a curve



Considered implementations which were decided against

Neural Network

Description

A neural network would have been constructed on the server side which would make use of backpropagation to shift weightings the in score-calculating equation to ensure that future trips would more closely approximate a value of 5.0.

Why this was not implemented

Neural networks, although demonstrably powerful, do not allow for mathematically sound statistical measurements of the accuracy of the algorithm. Because of their high complexity they are not easily shown to be useful (a core requirement stated by DVT).

Furthermore, neural networks would guarantee an average of around 5.0, but would not properly judge which variables best reflect good driving. It is very likely that a neural network may mistakenly place to great an emphasis on GPS coordinates or underestimate the effect of harsh acceleration on the quality of a person's driving.