Generating Random Numbers Using Social Media

This project investigates how random different random number generators are. A total of five random number generators were tested. Testing began November 3rd and lasted until January 10th. During these tests a script would continuously create random numbers and add them to a database. Each test collected three to four million random numbers. The purpose of this project was to see if better random numbers could be generated without the use of external hardware or a third-party generator service. The random numbers from each test were run through several different analyses. In the end a boxplot, and the p-values from a chi squared test on each were found. The chi squared p-value shows correlation in a set of data, the higher the p-value the more random the data is. The 4th version of the random number generator that used twitter to generate random numbers was the most random with a p-value of 56235223621678.4. My hypothesis stated that using photos from social media sites that were stored in a cache and randomly selected would yield the best random numbers. Unfortunately when building the random number generators I found that it is not practical to fetch photos from social media due to request limits set by such sites. It is more practical to use tweets to generate random numbers.