Is Google's 'Random Number Generator' Really Random?

By: Owen Erickson
3rd Hour AP Statistics
Mr. Selvaag
2017

Introduction

When you type certain things into Google's search engine, such as 'calculator', 'where am I?' or a translation question etc. a special result will appear, as I am sure you are all familiar with. When I was thinking about what I wanted my research question to be, I somehow came to searching 'Random Number Generator', which led to my discovery of Google's Random Number Generator. I knew some websites, primarily RANDOM.ORG, had their generator tested by statisticians and came to the conclusion that it was truly random, and I wanted to know if Google's version was also truly random.

Introduction: Research Question

I chose my research question to be “Is Google’s Random Number Generator really Random?” After some research, I found that randomness is not as binary as it seems, especially when it comes to computers, as there are many different programs, scripts and algorithms for many different coding languages. There is also the debate around whether or not ‘Pseudo-Random’ is random. Pseudo-Random refers to a method of generation in which a starting number is chosen, and there is a predetermined list of values corresponding to that seed. Random number tables, such as the one in our blue packets are an early version of this idea, with the line you start on being like a seed.

Data Collection: Method

To get my data, I just searched 'Random number generator' in Google and used the default range of 1-10 and clicked the 'GENERATE' button about five times and then put those values into a spreadsheet, then repeated this process until I had 1,000 values.

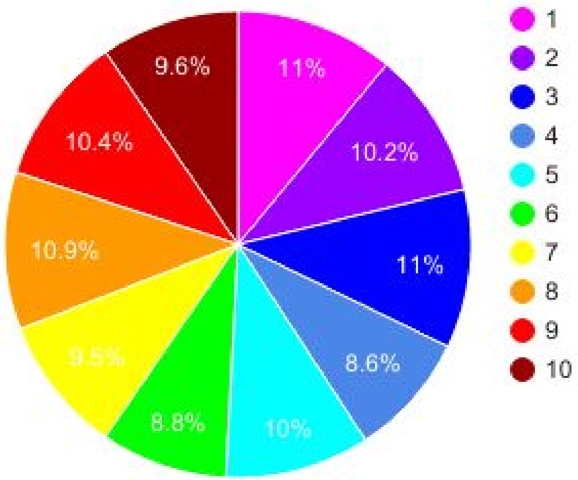
Data Collection: Analysis

Since randomness was the subject of my observational study, I did not have to take this into account when I was designing my data collection process. Similarly, I did not have to worry about independance, because I could have theoretically generated an infinite number of data values. The only parameter I had to meet with my data collection was to have a sample size larger than 50 to satisfy the Large Counts Condition.

Exploratory Data Analysis

Value	Observed Count	Expected Count
1	110	100
2	102	100
3	110	100
4	86	100
5	100	100
6	88	100
7	95	100
8	109	100
9	104	100
10	96	100

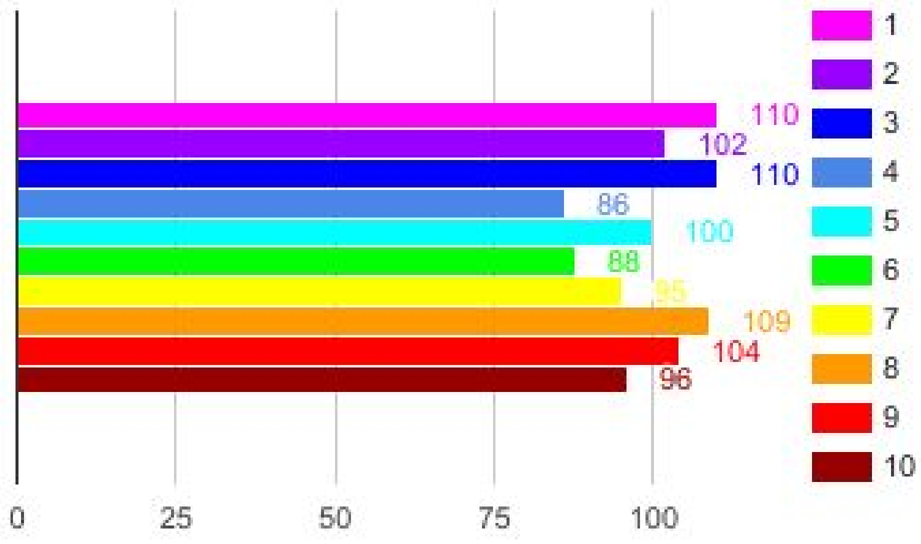
Exploratory Data Analysis



As you can see from this pie chart, there is very low deviation between values, with the largest being only 2.4%

Even though pie charts are typically only applicable to categorical data and mine is with numbers, it is considered categorical because the actual numbers assigned to each category are meaningless, they could be any 10 numbers.

Exploratory Data Analysis



Deviation can be seen slightly more clearly with this bar graph, but is still minimal.

'Fun' (meaningless) fact: the mean of my data was 5.449

Exploratory Data Analysis

Given the graphs shown, it is quite clear that there is not a significant bias or preference of the generator, and it appears that Google's Random Number Generator really is random.

Inference: Hypothesis for GOF Test

H_0: p_1=.1, p_2=.1, p_3=.1, p_4=.1, p_5=.1, p_6=.1, p_7=.1, p_8=.1, p_9=.1, p_{10}=.1

H_a: At least two of these proportions are incorrect.

Inference: Assumptions

Random: Given that this is what I was testing, I had to assume that it was indeed random to begin with, plus the fact that the program is claiming to be random, you could say that this condition was met via being stated in the problem.

Independant: There is a theoretically infinite population, meaning that regardless of my sample size, I could never take more than 10% of the population.

Large Counts Condition: In the probability distribution all expected Counts are 100, which is greater than 5.

Inference: Calculations for GOF Test

$X^2 = \sum(O-E)^2/E$ DoF=k-1 k=# of Kategories

$X^2 = \sum(110-100)^2/100, +(102-100)^2/100, +\ldots+(96-100)^2/100$ DoF=10-1

$X^2 = 6.82$ DoF=9

P=0.655856462

Inference: Conclusion

Because the P-value is not significant at the .05=α level, we fail to reject the null hypothesis that the generator is random.

Summary: Answering Research Question

In my Research Question, I asked "Is Google's Random Number Generator really Random?" Given the data gathered and the calculated P-value, it is very safe to say; "Yes, Google's Random Number Generator really is Random."

Conclusion: Improvements & Limitations

Throughout my data collection process, I had to close the tabs in which I was obtaining my data, which could have resulted in generating new seeds, but this only applies if the generator is Pseudo-Random. Given that I documented 1000 numbers in sets of five at a time, it is very possible that I misremembered or mistyped some values into my spreadsheet.

If I were to repeat this experiment, I would increase the sample size and use a range of 0-9 instead of 1-10 just in case having a two-digit number affected the data somehow.

Thanks for Listening!