HUMAN PERCEPTION OF RANDOMNESS IN NUMBER GENERATION

IESM315 - Design and Analysis of Experiments

American University of Armenia

A. Gumruyan | I. Torosyan

MOTIVATION

01

Understanding the nature of human-generated random sequences

02

Investigating influencing factors and identifying patterns

03

The potential of practical implications for decision-making

LITERATURE REVIEW

REFLECTING ON SURVEY COMMENTS

Perception of Randomness

“I think they are not that random because there is some kind of a reason why my brain thought of that exact numbers and not other ones.”

“I recalled number bias and it would be interesting if a thing like that really is a thing”

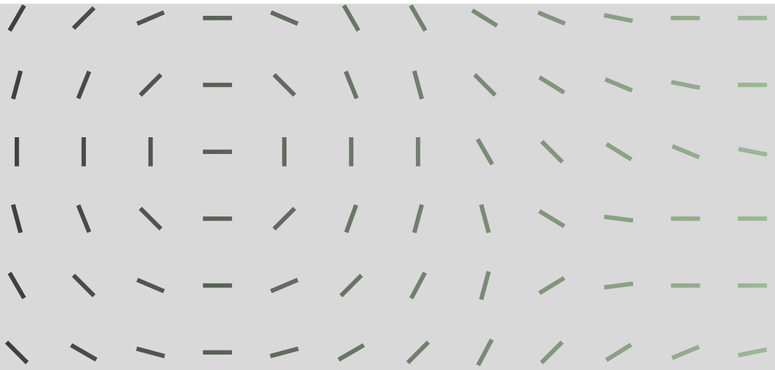
How Were The Numbers Picked

“I used random num generator”

“Pen-and-paper method for generating random numbers by John von Neumann”

“I have been thinking about my fav numbers and it was not easy to let them go to be more "random"”

> **Predictable cognitive biases influence our choices, and seemingly random decisions follow discernible patterns.**



— Predictably Irrational

Dan Ariely

> When faced with tasks involving randomness or uncertainty, people often rely on mental shortcuts, or heuristics, that lead to predictable decision-making patterns.

— Thinking, Fast and Slow

Daniel Kahneman

> People tend to overweight the likelihood of rare events when judging them based on descriptions.
>
> When overwhelmed with information, we resort to simpler heuristics.

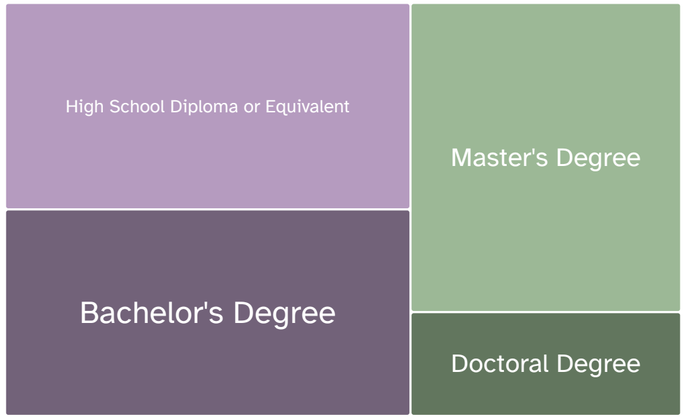
— The Description-Experience Gap in Rare Events

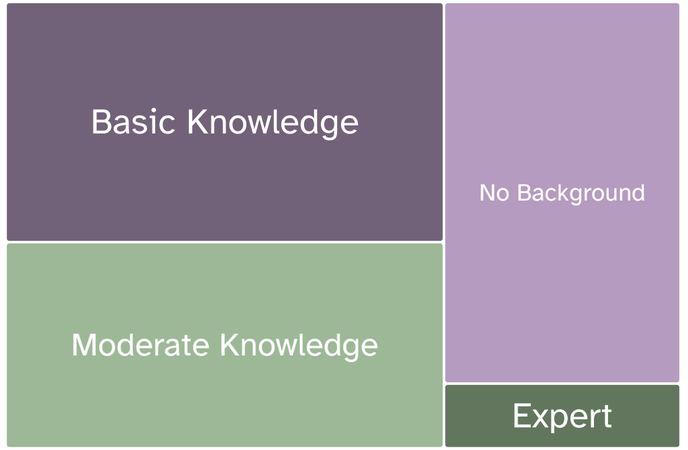
Aydogan and Gao

METHODOLOGY

DATA COLLECTION

Method	**Google Forms Survey**
Questionnaire	**Informed Consent** **Education** 1. Current/Most Recent Degree 2. Statistics Background **Random Number Sequence Generation** 1. Unrestricted 2. Restricted: from 1 to 50 3. Restricted: above 70 **Personal Perception** 1. Randomness Assessment 2. Additional Comments
Status	Nov 10th to Nov 23rd, 2023 **Participants:** 100





DATA MANIPULATION

Stage 1: Naked Eye Scanning	**Identification and exclusion of the creative participant who provided numbers beyond representable integer limits** (e.g. 10^124, while the limit is 2*10^9)
Stage 2: Exploratory Data Analysis	**Identification of major outlier occurrences and subsequent filtering following the Tukey Method of outlier labeling** (keeping values in the range {Q1 – 1.5*IQR), Q3 + 1.5*IQR})

Figure 1: Distribution of random number sequences before outlier adjustment (order: No Restriction, Restricted from 1 to 50, Restricted above 70)

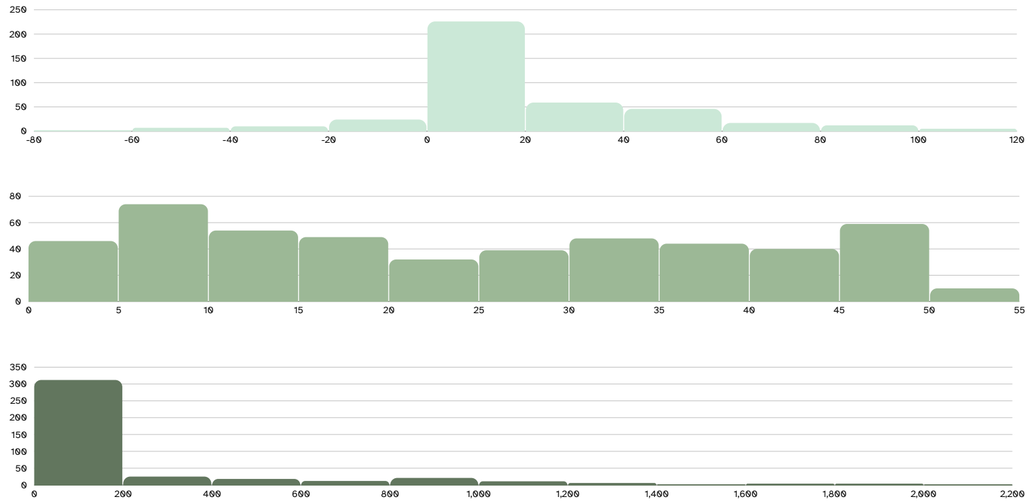


Figure 2: Distribution of random number sequences after outlier adjustment (order: No Restriction, Restricted from 1 to 50, Restricted above 70)

DATA ANALYSIS

Randomness

Runs Test to assess the randomness of each generated sequence.

Chi-Squared Test to compare the observed frequency of numbers to the expected frequency in a truly random sequence.

Density Distribution

Kolmogorov-Smirnov Test *to determine if the generated sequences follow known distributions (Normal, Uniform, Poisson, or Exponential).*

Kernel-smoothed Density Plots for the visual aspect of analysis.

Group Comparison

One-way ANOVA to examine if there are significant differences among the participant groups based on statistics background.

Ad hoc Tukey T-test to assess the significance of differences between all pairs of statistics background groups.

Correlation

Spearman's Rank Correlation Test *to assess the monotonic relationship between participants' statistics background and randomness assessment in random number generation.*

RESULTS & CONCLUSION

RANDOMNESS

Runs Test

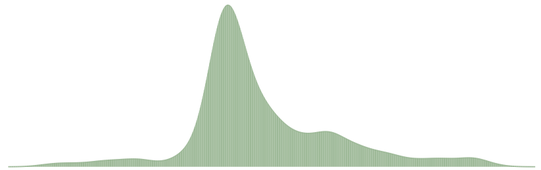
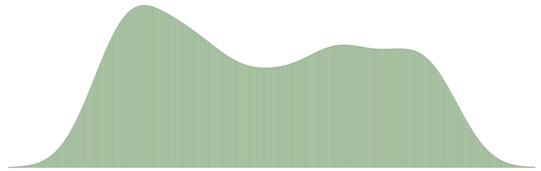
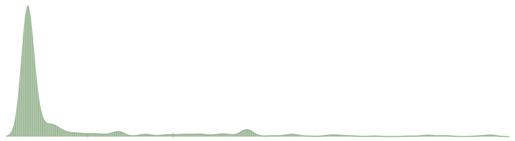
Unrestricted	Standardized Runs Statistic: -6.0473 P-value: 1.473e-09 ***Significant deviation from randomness.***
Restricted from 1 to 50	Standardized Runs Statistic: -1.7405 P-value: 0.08176 ***Participants performed reasonably well, achieving relative randomness.***
Restricted above 70	Standardized Runs Statistic: -3.8732 P-value: 0.0001074 ***Significant deviation from randomness.***

Chi-Squared Test

Unrestricted	Chi-Squared Statistic: 632.08 P-value: < 2.2e-16 ***Significant deviation from random frequency distribution.***
Restricted from 1 to 50	Chi-Squared Statistic: 235.8 P-value: < 2.2e-16 ***Significant deviation, yet a relatively smaller chi-squared statistics.***
Restricted above 70	Chi-Squared Statistic: 892.04 P-value: < 2.2e-16 ***Significant deviation from random frequency distribution.***

DENSITY DISTRIBUTION

Kolmogorov-Smirnov Test | Kernel-smoothed Density Plots

	Exponential	Uniform	Poisson	Normal	
Unrestricted	D = 0.15503 p-value = 6.078e-09	D = 1 p-value < 2.2e-16	D = 0.41062 p-value < 2.2e-16	D = 0.058148 p-value = 0.1267	
Restricted: 1 to 50	D = 0.15475 p-value = 1.012e-10	D = 1 p-value < 2.2e-16	D = 0.41758 p-value < 2.2e-16	D = 0.057982 p-value = 0.0717	
Restricted: above 70	D = 0.15499 p-value = 3.979e-09	D = 1 p-value < 2.2e-16	D = 0.41144 p-value < 2.2e-16	D = 0.058127 p-value = 0.1194	

GROUP COMPARISON

ANOVA Test

Unrestricted	F-value: 2.972 P-value: 0.0317 ***Difference of means in at least one level of statistics background.***
Restricted from 1 to 50	F-value: 1.664 P-value: 0.174 ***No difference in means across levels of statistics background.***
Restricted above 70	F-value: 2.575 P-value: 0.0535 ***Marginal difference in means, above the conventional significance level.***

Ad Hoc Tukey T-test

Unrestricted	***Significant difference between participants with No Background and Moderate Knowledge.***
Restricted above 70	***Marginal significance in the difference between participants with No Background and Experts, not reaching the conventional significance level.***

CORRELATION

Spearman’s Rank Correlation Test

Unrestricted	Weak negative monotonic correlation (-0.12). P-value: 0.2547. ***Not statistically significant. No strong correlation observed between statistics background and randomness assessment.***
Restricted from 1 to 50	Weak negative monotonic correlation (-0.06). P-value: 0.5466. ***Not statistically significant. No strong correlation found between statistics background and randomness in the specified range.***
Restricted above 70	Very weak positive monotonic correlation (0.02). P-value: 0.8163. ***Not statistically significant. No substantial correlation identified between statistics background and randomness assessment for sequences above the threshold.***

CONCLUSION

01 Individuals face challenges in producing truly random sequences, particularly when constraints are imposed.

02 Statistical backgrounds play a role in random number generation but is not a decisive factor on its own.

THANK YOUR FOR YOUR ATTENTION!

Questions? Concerns?