**Detecting Human Bias in Random Number Generation Using Probability and Entropy**

**Author’s Note**My name is Rithwik Garnaik, a rising sophomore from North Carolina with a passion for mathematics, statistics, and understanding how people think. I've always been fascinated by how predictable human behavior can be, even when we try to be completely random. This independent research project explores a mathematical approach to studying human decision-making, particularly our inability to create truly random sequences.

**Abstract**This study investigates the biases that appear when humans attempt to generate random sequences without the aid of a randomizing device. By collecting human-generated sequences of binary choices and comparing them to computer-generated randomness, this research identifies patterns in human choices. Statistical analysis and entropy measurement are used to quantify how predictable human "randomness" really is. The findings suggest that humans tend to over-alternate, avoid repetition, and produce lower-entropy sequences than true randomness would suggest.

**Introduction**While randomness is a core idea in probability and statistics, humans have been shown to be surprisingly bad at generating random outcomes. This phenomenon has implications in psychology, computer science, and cryptography. This paper aims to model and detect patterns in human-generated sequences using basic statistical analysis and entropy, a concept from information theory that measures unpredictability.

**Research Question**Can we detect and model human biases in self-generated random sequences using statistical tools like probability and entropy?

**Methodology**

**Data Collection**Participants (friends, classmates, and family) were asked to:

* Flip a coin 50 times and record the sequence (or pretend to and generate what they believe is random)
* OR simply write down a sequence of 50 Hs and Ts that they believe mimics coin flips

A separate set of sequences was generated using a Python random number generator or a random.org tool to serve as a control group.

**Data Analysis**Each sequence was analyzed based on:

* **Frequency Distribution**: Number of heads vs. tails.
* **Alternation Rate**: How often the result changes (H to T or T to H).
* **Repetition Gaps**: Lengths of repeated symbols.
* **Entropy**: Using the Shannon entropy formula: Entropy near 1.0 is expected in a fair coin flip; lower entropy suggests predictability.

**Results**The human-generated sequences showed:

* **Higher alternation rates** (~70%) compared to the expected 50% from true randomness.
* **Fewer long runs** (e.g., 4+ Hs or Ts in a row were rare in human sequences).
* **Lower entropy values**, averaging 0.83, compared to 0.99 from computer-generated sequences.

These results confirm a clear human bias toward alternating outcomes and avoiding repetition. Participants often subconsciously tried to "balance" the sequence, thinking that randomness should look more evenly mixed than it really does.

**Discussion**The patterns found in human sequences are consistent with prior psychological findings. People equate randomness with uniformity and thus overcompensate by switching choices more than random chance would dictate. This misunderstanding is known as the **gambler's fallacy**, where people expect outcomes to "even out" in the short term.

Understanding this bias has real-world applications:

* In security, where passwords or security keys should be random, human biases can weaken systems.
* In experimental design, assuming unbiased human behavior in randomization can lead to flawed results.
* In teaching probability, this serves as a powerful example of intuition vs. math.

**Limitations**

* Small sample size (~20 participants)
* No control for age or mathematical background
* Only binary sequences (not digits 0–9 or more complex choices)

**Conclusion**This research shows that humans are not good at creating truly random sequences. With basic probability tools and entropy analysis, we can identify and measure the predictable biases in how we generate "random" choices. This study not only highlights a gap between perception and reality but also demonstrates how mathematical thinking can be applied to human behavior in everyday situations.

**Possible Extensions**

* Analyze longer sequences or non-binary options (e.g., 0-9 digits)
* Compare sequences by age, background, or familiarity with statistics
* Use AI to detect and predict the next choice in a human sequence