Let's try to model English text as a Markov chain with states $\{a, b, c, d, \ldots, z, \text{space}\}$ (we ignore numbers and punctuation).

1 Estimate the transition probabilities $p(q, u)$, $p(\text{space}, u)$ and $p(u, \text{space})$.

· · · · · ·

2 Estimate the probability that the first 6 letters spell out "markov", given that $X_0 = m$.

· · · · · ·

3 Estimate $\pi(a)$, $\pi(z)$ and $\pi(\text{space})$. Describe applications to Scrabble.

· · · · · ·

4 Estimate $E_{\text{space}}(T_{\text{space}})$.

· · · · · ·

5 How would you estimate the transition probabilities from real-world data?

· · · · · ·

6 Write down a plausible sequence $(X_0, X_1, \ldots, X_{10})$. Does it look like English?

· · · · · ·

7 Invent a more accurate model of English text. (It might or might not be a Markov chain!)

· · · · · ·