Nucleus Sampling

P(y | they live in a remote desert by)

0.01	roads
0.01	towns
0.01	people
0.005	streets

→ renormalize and sample

~~~~~~~~~~ cut off after p% of probability mass

- Define a threshold p. keep the most probable options accounting for p% of probability mass (the nucleus), then sample among these.

- To implement: Sort options by probability, truncate the list once the total exceeds p, then renormalize and sample from it.

* Decent perplexity, doesn't have bad repititions like greedy/beam search
~~~~~~~~~~