Central Question in LM: $p(w_1w_2…w_m|C)=?$

- What is C? We usually ignore C (="context") since it depends on the application, but it's important to consider it when applying a LM
  - Refinement 1: $p(w_1w_2…w_m|C) \approx p(w_1w_2…w_m)$
- What random variables are involved? What is the event space?
  - What event does "$w_1w_2…w_m$" represent? What is the sample space?
  - $p(w_1w_2…w_m) = p(X=w_1w_2…w_m)$ vs. $p(X_1=w_1,X_2=w_2,…X_m=w_m)$?

X=?

X_1=? **X_2=?** **X_m=?**