$$\text{perplexity}(W) = P(w_1 w_2 \dots w_N)^{-\frac{1}{N}} \tag{3.14}$$
$$= \sqrt[N]{\frac{1}{P(w_1 w_2 \dots w_N)}}$$

Or we can use the chain rule to expand the probability of W:

$$\text{perplexity}(W) = \sqrt[N]{\prod_{i=1}^{N} \frac{1}{P(w_i | w_1 \dots w_{i-1})}} \tag{3.15}$$

$$\text{perplexity}(W) = \sqrt[N]{\frac{1}{P(w_1 w_2 \cdots w_n)}}$$

由概率的链式法则得

$$P(w_1 w_2 \cdots w_n) = P(w_2|w_1) \cdot P(w_3|w_1 w_2) \cdots P(w_i | w_1 \cdots w_{i-1})$$

即 $$P(w_1 w_2 \cdots w_n) = \prod_{i=1}^{N} P(w_i | w_1 \cdots w_{i-1})$$

代入 perplexity(W)得

$$\text{perplexity}(W) = \sqrt[N]{\prod_{i=1}^{N} \frac{1}{P(w_i | w_1 \cdots w_{i-1})}}$$

得证。