



$$w_1 \ w_2 \ \cdots \ w_{t-1} \ \boxed{w_t} \ \cdots \ w_m$$

$$= P(w_t | w_1, w_2, \dots, w_{t-1})$$

図 5-3 言語モデルが扱う事後確率： t 番目の単語をターゲットとして、 t 番目より左側の単語すべてをコンテキスト（条件）とする