

To find the probability p_i of token i (not in top- k tokens), add bias λ to the token’s logit ℓ_i and obtain the biased logprob

$$\log p'_i = \log \frac{\exp(\ell_i + \lambda)}{\exp(\ell_i + \lambda) + \sum_{j \neq i} \exp \ell_j}. \quad (1)$$

Rearranging this, we obtain

$$p'_i = \frac{\exp(\ell_i + \lambda)}{\exp(\ell_i + \lambda) + \sum_{j \neq i} \exp \ell_j} \quad (2)$$

$$p'_i \left(\exp(\ell_i + \lambda) + \sum_{j \neq i} \exp \ell_j \right) = \exp(\ell_i + \lambda) \quad (3)$$

$$p'_i \exp(\ell_i + \lambda) + p'_i \sum_{j \neq i} \exp \ell_j = \exp(\ell_i + \lambda) \quad (4)$$

$$\exp(\ell_i + \lambda) - p'_i \exp(\ell_i + \lambda) = p'_i \sum_{j \neq i} \exp \ell_j \quad (5)$$

$$(1 - p'_i) \exp(\ell_i + \lambda) = p'_i \sum_{j \neq i} \exp \ell_j \quad (6)$$

$$\exp \lambda \exp \ell_i \frac{1 - p'_i}{p'_i} = \sum_{j \neq i} \exp \ell_j. \quad (7)$$

Note that the righthand side does not depend on the bias term, therefore we can equate the lefthand side to the the case where $\lambda = 0$ (i.e., the unbiased case) and solve for the probability p_i

$$\exp \lambda \exp \ell_i \frac{1 - p'_i}{p'_i} = \exp 0 \exp \ell_i \frac{1 - p_i}{p_i} \quad (8)$$

$$\exp \lambda \frac{1 - p'_i}{p'_i} = \frac{1 - p_i}{p_i} \quad (9)$$

$$p_i = \frac{1}{\exp \lambda \frac{1 - p'_i}{p'_i} + 1}. \quad (10)$$

Thus, it is possible to obtain unbiased logprobs for any token with exactly 1 API call.