

# Temperature Calculation: Step-by-Step

Given: Logits = [2.0, 1.0, 0.5, 0.2]  
Tokens = ["cat", "dog", "bird", "fish"]

**Step 1: Scale Scaled by 0.5 (MORE FOCUSED)**  
Scaled =  $\frac{[2.0, 1.0, 0.5, 0.2]}{\max(2.0, 1.0, 0.5, 0.2) \cdot 0.5} = [0.6, 0.5, 0.19, 0.12]$   
Softmax → 73% on "cat" (BALANCED)  
→ 73% on "cat" (VERY FOCUSED)

**Step 3: Scale by T=2.0 (FLATTER)**  
Scaled Softm  $p_i = \frac{\exp(\text{logit}_i/T)}{\sum_j \exp(\text{logit}_j/T)}$   
CH FLATTER)

Lower  $T \rightarrow$  More confident (peaky)  
Higher  $T \rightarrow$  More random (flat)