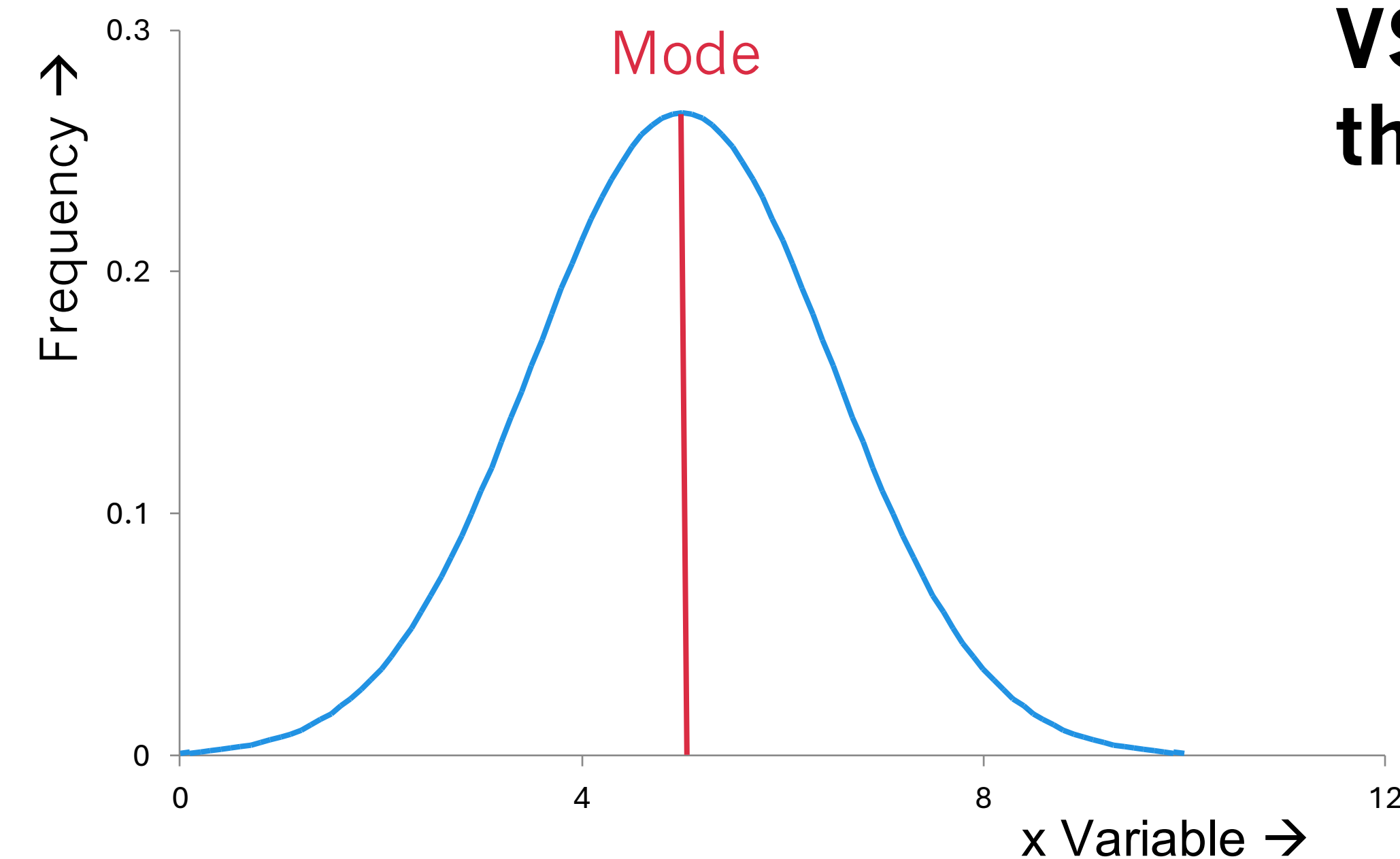# Proof: RLHF Causes Mode Collapse

Step 1: Diverse pre-training data → Base Model

Step 2: Base Model after RLHF → Mode Collapse

VS recovers the diversity

Solution: Verbalize Sampling (VS)

Direct Prompting

User Instruction

Tell a joke about coffee.

X 5

Model Outputs

Why did the coffee file a police report? Because it got mugged! ☕😊
output x1

Why did the coffee file a police report? Because it got mugged! ☕😊
output x2

Why did the coffee file a police report? Because it got mugged! ☕😊
output x3

Why did the coffee file a police report? Because it got mugged! ☕😊
output x4

Why did the coffee file a police report? Because it got mugged! ☕😊
output x5

Verbalized Sampling Prompting

User Instruction

Generate 5 responses to the user instruction, and their corresponding probabilities:
Tell a joke about coffee.

Model Outputs

How does a coffee show affection? It gives you a latte love! *(Prob: 0.18)*

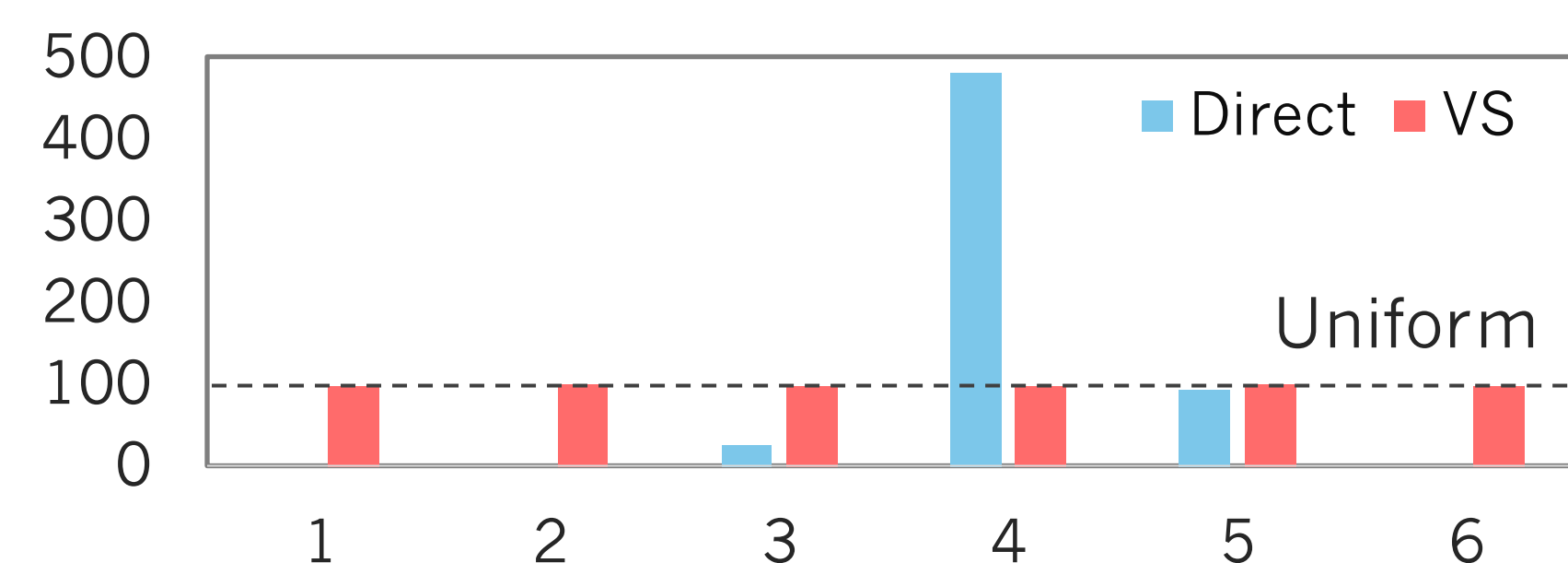What do you call a coffee that's been sitting out too long? A has-bean! *(Prob: 0.12)*

Why did the latte go to therapy? It had too much foam to deal with. *(Prob: 0.15)*

What do you call sad coffee? Depresso. *(Prob: 0.10)*

What's a coffee's favorite spell? Espresso Patronum! *(Prob: 0.06)*

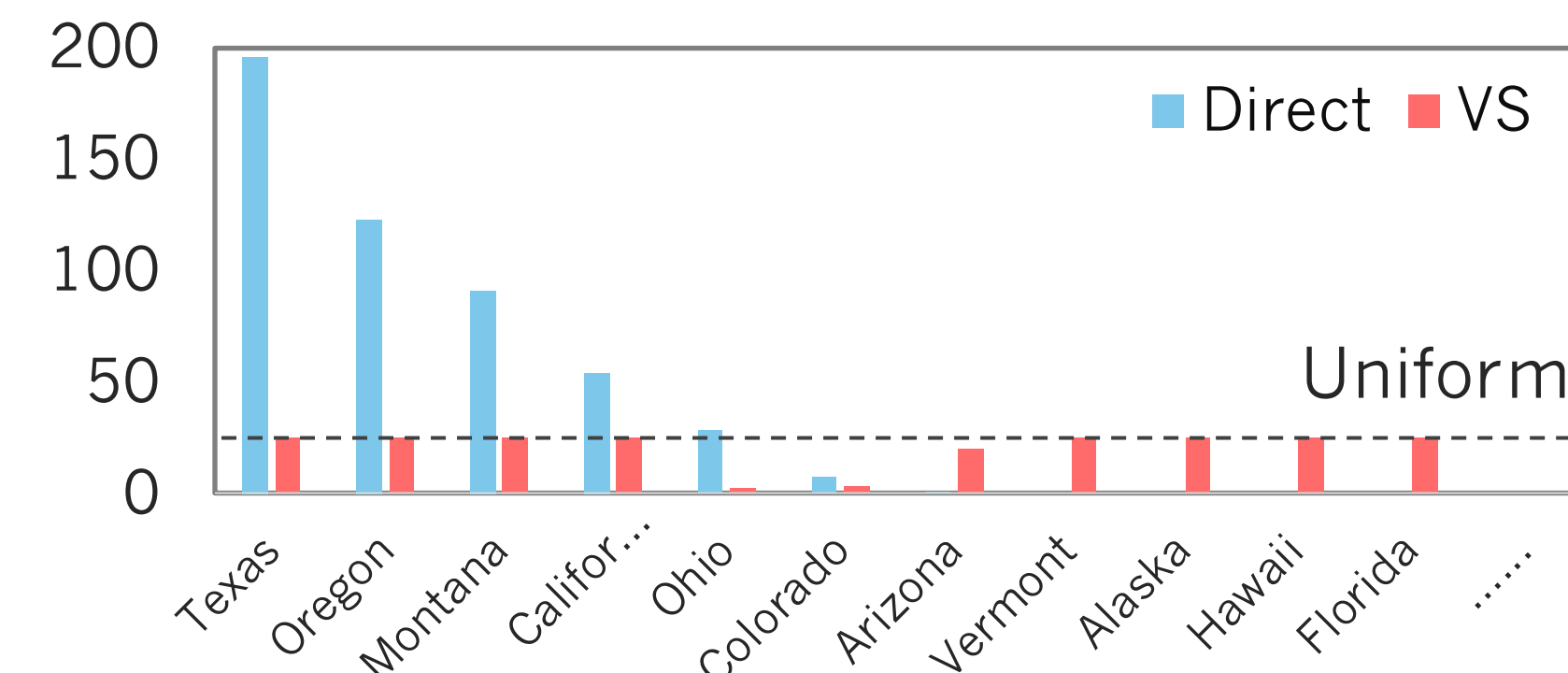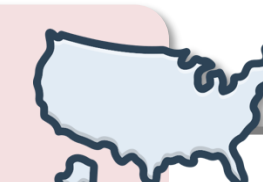# Application: Effective On Variety of Tasks

Random Number Generation 🎲

Instruction: Roll a fair six-sided dice.

Output (VS):
- 1 *(Prob: 0.167)*
- 2 *(Prob: 0.167)*
- 3 *(Prob: 0.167)*
- 4 *(Prob: 0.167)*
- 5 *(Prob: 0.167)*
- 6 *(Prob: 0.167)*
- .....

State Name Generation

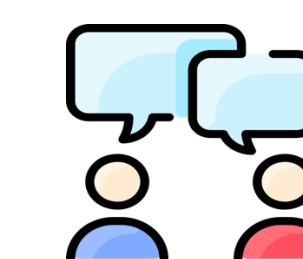**Instruction:** Name a US state.

Output (VS):
- California *(Prob: 0.25)*,
- Alaska *(Prob: 0.17)*,
- Vermont *(Prob: 0.12)*,
- Maine *(Prob: 0.15)*,
- South Dakota *(Prob: 0.11)*,
- ......

Story Writing

Poem Writing

Dialogue Simulation

Synthetic negative data generation

......