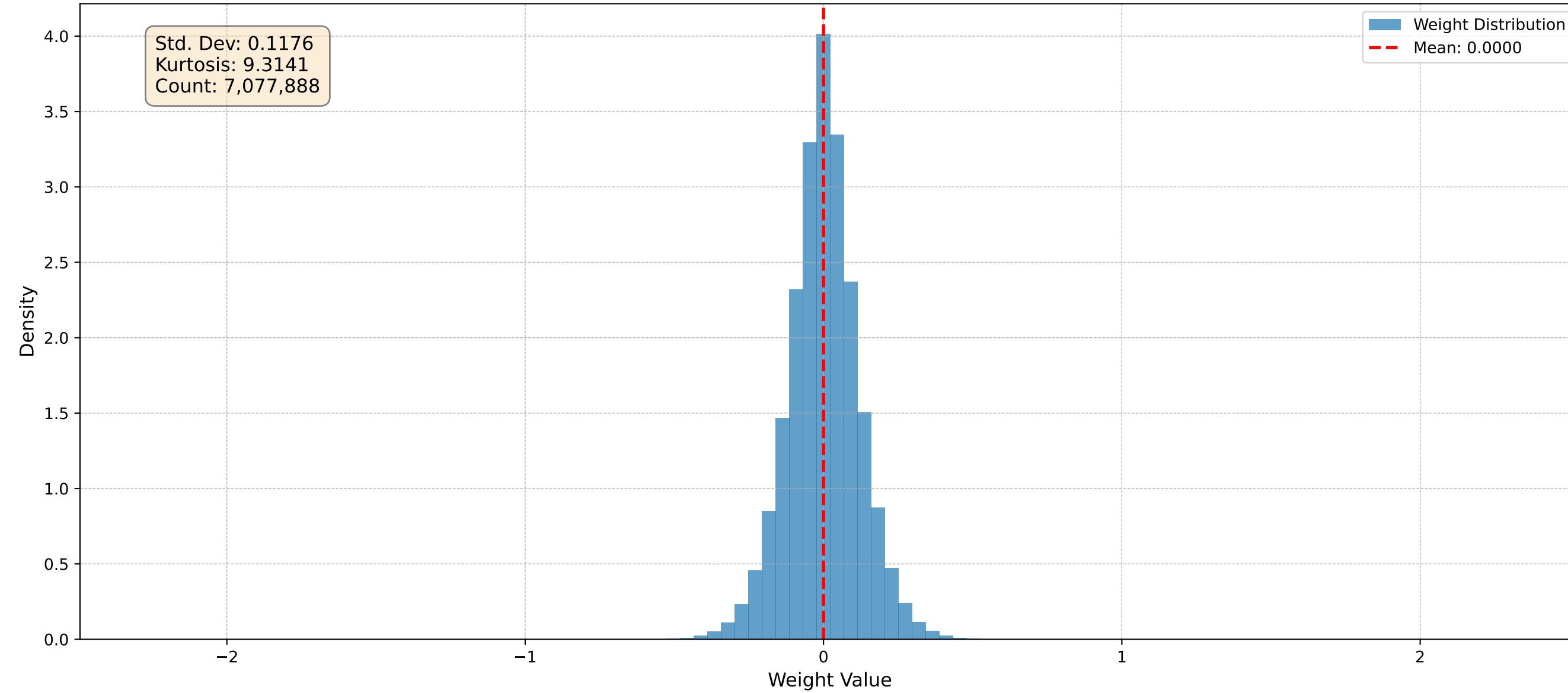Distributions for "attention_value" in gpt2