cos similarity histogram – layer 20
bert-large-uncased_uniform-tokens_⟨Qx,Ky⟩_trained-params_dense-on_24-laye