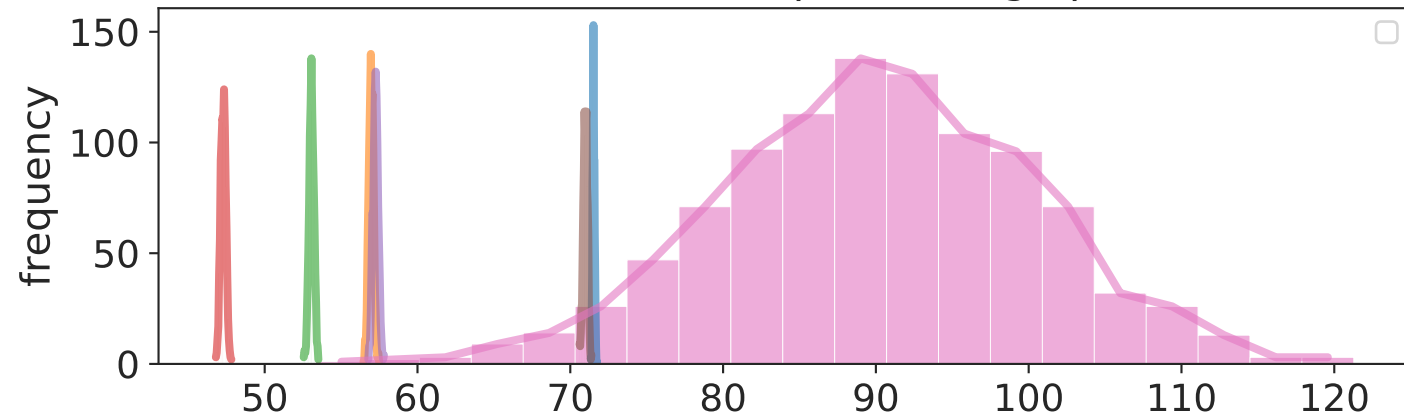
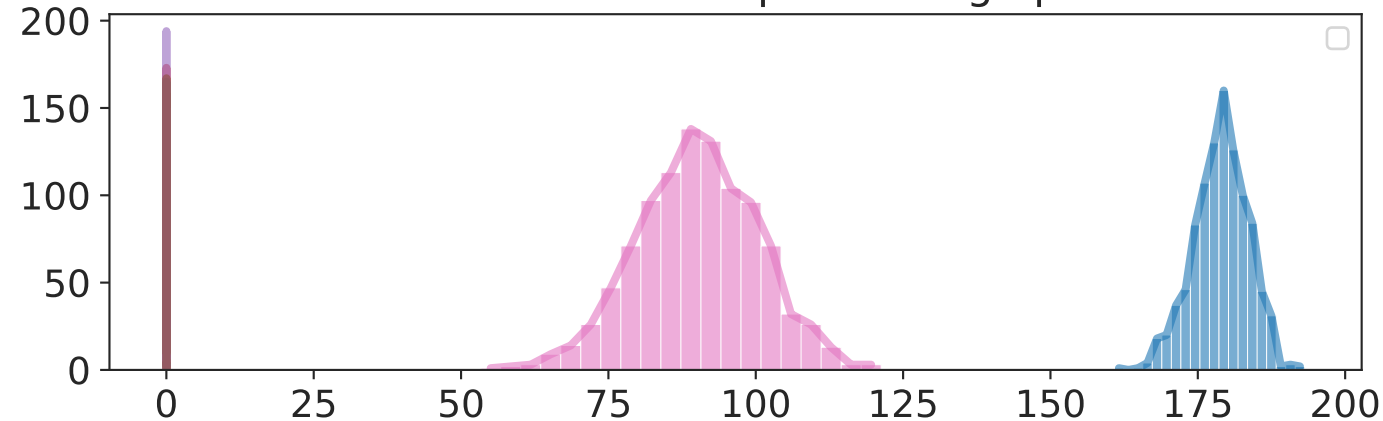


# Generated distributions when increasing pretraining epochs

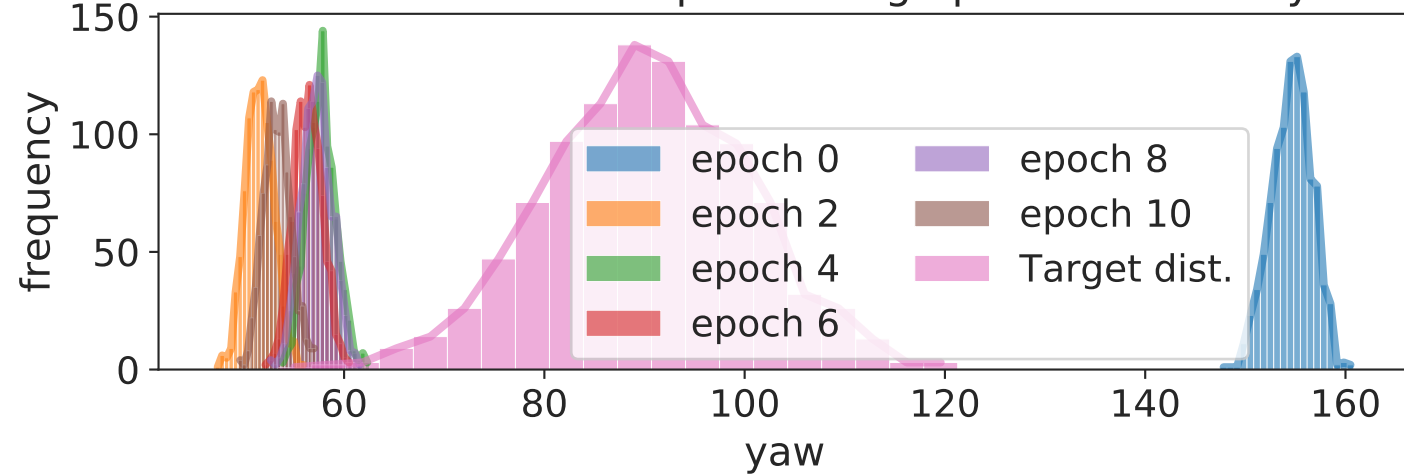
Distribution after 5 pretraining epochs



Distribution after 10 pretraining epochs



Distribution after 5 pretraining epochs + lr decay



Distribution after 10 pretraining epochs + lr decay

