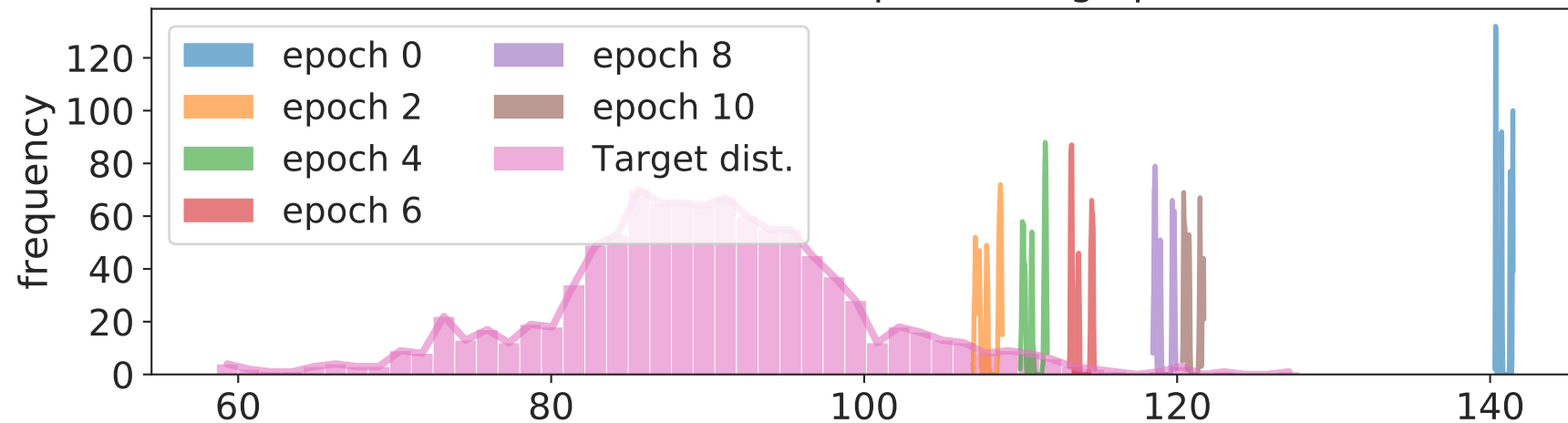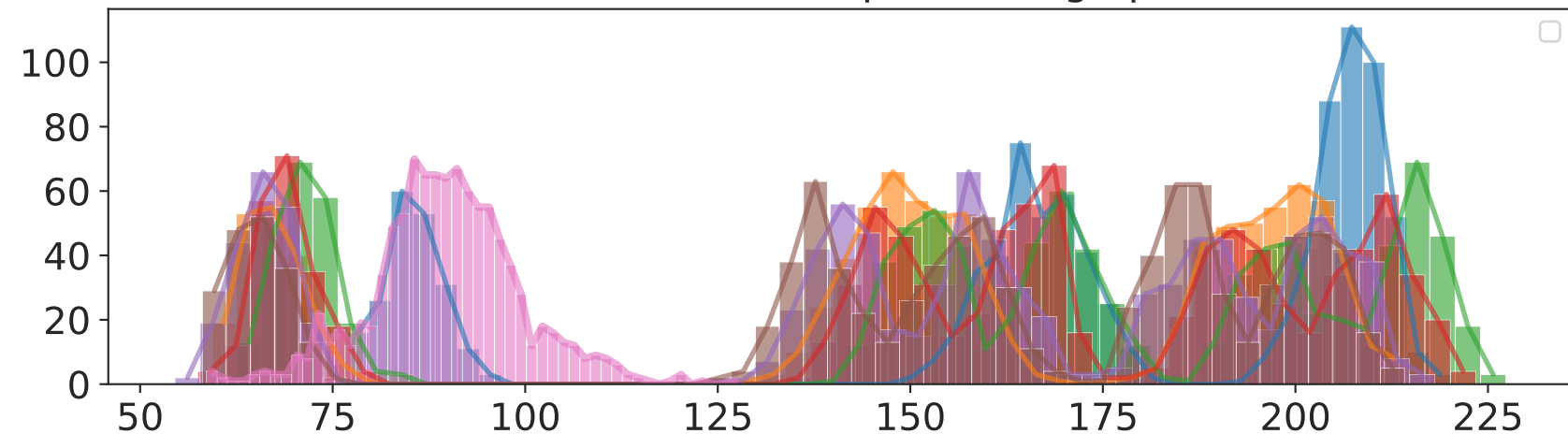**Generated distributions for multi-class 1 modifiable attribute experiments**
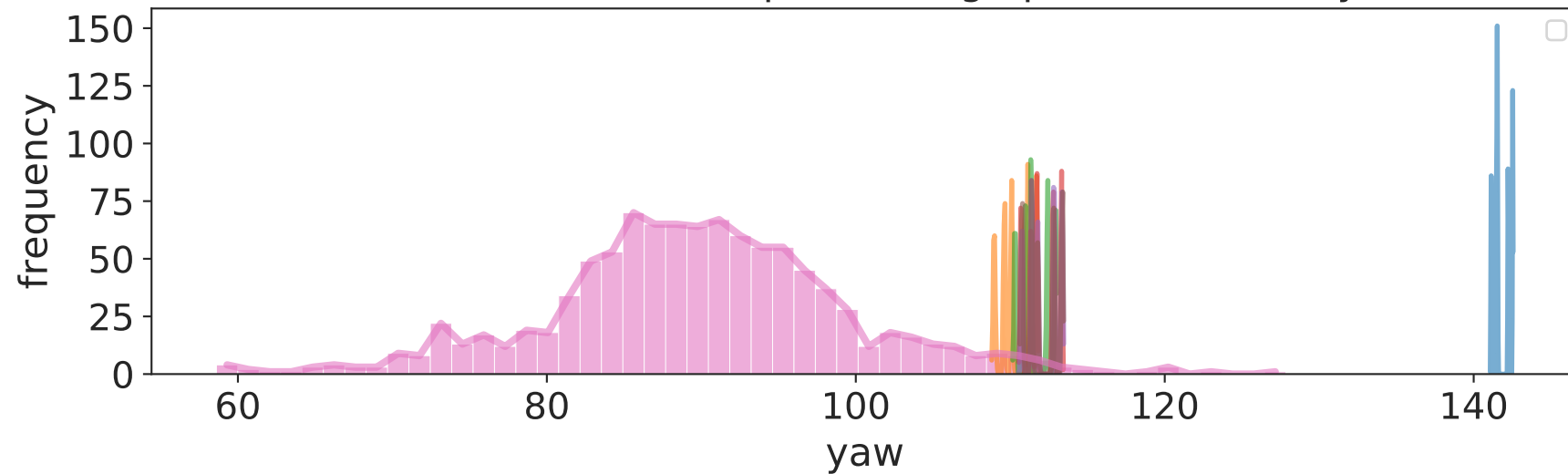
Distribution after 1 pretraining epochs

Distribution after 10 pretraining epochs

- epoch 0
- epoch 2
- epoch 4
- epoch 6
- epoch 8
- epoch 10
- Target dist.

Distribution after 1 pretraining epochs + lr decay

Distribution after 10 pretraining epochs + lr decay