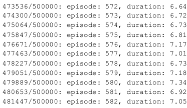
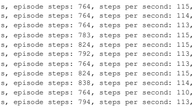
```
Training for 500000 steps ...
    913/500000: episode: 1, duration: 3.370s, episode steps: 913, steps per second: 271, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 2.413 [0.000, 5.000], mean observation: 105.605 [87.000, 236.000], loss: --, mean_absolute_error: --, mean_q: --, mean_eps: --
   1733/500000: episode: 2, duration: 2.540s, episode steps: 820, steps per second: 323, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 2.420 [0.000, 5.000], mean observation: 105.615 [87.000, 236.000], loss: --, mean_absolute_error: --, mean_q: --, mean_eps: --
   2755/500000: episode: 3, duration: 3.143s, episode steps: 1022, steps per second: 325, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean action: 2.391 [0.000, 5.000], mean observation: 105.615 [87.000, 236.000], loss: --, mean_absolute_error: --, mean_q: --, mean_eps: --
   3658/500000: episode: 4, duration: 2.832s, episode steps: 903, steps per second: 319, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 2.501 [0.000, 5.000], mean observation: 105.444 [87.000, 236.000], loss: --, mean_absolute_error: --, mean_q: --, mean_eps: --
   4754/500000: episode: 5, duration: 3.388s, episode steps: 1096, steps per second: 324, episode reward: -20.000, mean reward: -0.018 [-1.000, 1.000], mean action: 2.451 [0.000, 5.000], mean observation: 105.562 [87.000, 236.000], loss: --, mean_absolute_error: --, mean_q: --, mean_eps: --
   5606/500000: episode: 6, duration: 2.638s, episode steps: 852, steps per second: 323, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.446 [0.000, 5.000], mean observation: 105.598 [87.000, 236.000], loss: --, mean_absolute_error: --, mean_q: --, mean_eps: --
   6819/500000: episode: 7, duration: 3.826s, episode steps: 1213, steps per second: 317, episode reward: -18.000, mean reward: -0.015 [-1.000, 1.000], mean action: 2.342 [0.000, 5.000], mean observation: 105.513 [87.000, 236.000], loss: --, mean_absolute_error: --, mean_q: --, mean_eps: --
   7847/500000: episode: 8, duration: 3.197s, episode steps: 1028, steps per second: 322, episode reward: -21.000, mean reward: -0.020 [-1.000, 0.000], mean action: 2.250 [0.000, 5.000], mean observation: 105.605 [87.000, 236.000], loss: --, mean_absolute_error: --, mean_q: --, mean_eps: --
   8733/500000: episode: 9, duration: 2.783s, episode steps: 886, steps per second: 318, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.270 [0.000, 5.000], mean observation: 105.626 [87.000, 236.000], loss: --, mean_absolute_error: --, mean_q: --, mean_eps: --
   9637/500000: episode: 10, duration: 2.794s, episode steps: 904, steps per second: 324, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 2.287 [0.000, 5.000], mean observation: 105.609 [87.000, 236.000], loss: --, mean_absolute_error: --, mean_q: --, mean_eps: --
  10709/500000: episode: 11, duration: 9.092s, episode steps: 1072, steps per second: 118, episode reward: -19.000, mean reward: -0.018 [-1.000, 1.000], mean action: 2.448 [0.000, 5.000], mean observation: 105.493 [87.000, 236.000], loss: 0.012008, mean_absolute_error: 0.024514, mean_q: 0.006129, mean_eps: 0.906796
  11789/500000: episode: 12, duration: 9.550s, episode steps: 1080, steps per second: 113, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action: 2.590 [0.000, 5.000], mean observation: 105.525 [87.000, 236.000], loss: 0.011468, mean_absolute_error: 0.024827, mean_q: 0.008433, mean_eps: 0.898750
  12669/500000: episode: 13, duration: 7.810s, episode steps: 880, steps per second: 113, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.470 [0.000, 5.000], mean observation: 105.621 [87.000, 236.000], loss: 0.011236, mean_absolute_error: 0.017738, mean_q: 0.009588, mean_eps: 0.889930
  13540/500000: episode: 14, duration: 7.651s, episode steps: 871, steps per second: 114, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.524 [0.000, 5.000], mean observation: 105.608 [87.000, 236.000], loss: 0.011932, mean_absolute_error: 0.021682, mean_q: 0.013629, mean_eps: 0.882064
  14477/500000: episode: 15, duration: 8.266s, episode steps: 937, steps per second: 113, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action: 2.633 [0.000, 5.000], mean observation: 105.519 [87.000, 236.000], loss: 0.011370, mean_absolute_error: 0.018672, mean_q: 0.001063, mean_eps: 0.873928
  15408/500000: episode: 16, duration: 8.164s, episode steps: 931, steps per second: 114, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 2.528 [0.000, 5.000], mean observation: 105.475 [87.000, 236.000], loss: 0.011930, mean_absolute_error: 0.032536, mean_q: 0.004579, mean_eps: 0.865522
  16250/500000: episode: 17, duration: 7.362s, episode steps: 842, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.393 [0.000, 5.000], mean observation: 105.465 [87.000, 236.000], loss: 0.011137, mean_absolute_error: 0.026250, mean_q: 0.019854, mean_eps: 0.857548
  17232/500000: episode: 18, duration: 8.701s, episode steps: 982, steps per second: 113, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean action: 2.720 [0.000, 5.000], mean observation: 105.593 [87.000, 236.000], loss: 0.013198, mean_absolute_error: 0.017801, mean_q: -0.004130, mean_eps: 0.849340
  17996/500000: episode: 19, duration: 6.814s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.795 [0.000, 5.000], mean observation: 105.595 [87.000, 236.000], loss: 0.011465, mean_absolute_error: 0.041234, mean_q: -0.034546, mean_eps: 0.841492
  18760/500000: episode: 20, duration: 6.651s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.772 [0.000, 5.000], mean observation: 105.595 [87.000, 236.000], loss: 0.010882, mean_absolute_error: 0.053741, mean_q: -0.048311, mean_eps: 0.834616
  19602/500000: episode: 21, duration: 7.433s, episode steps: 842, steps per second: 113, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.894 [0.000, 5.000], mean observation: 105.423 [87.000, 236.000], loss: 0.011817, mean_absolute_error: 0.046723, mean_q: -0.039065, mean_eps: 0.827380
  20681/500000: episode: 22, duration: 9.439s, episode steps: 1079, steps per second: 114, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action: 2.080 [0.000, 5.000], mean observation: 105.497 [87.000, 236.000], loss: 0.010678, mean_absolute_error: 0.016258, mean_q: 0.003136, mean_eps: 0.818722
  21583/500000: episode: 23, duration: 8.030s, episode steps: 902, steps per second: 112, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 2.725 [0.000, 5.000], mean observation: 105.461 [87.000, 236.000], loss: 0.012358, mean_absolute_error: 0.010748, mean_q: -0.001192, mean_eps: 0.809812
  22621/500000: episode: 24, duration: 9.283s, episode steps: 1038, steps per second: 112, episode reward: -19.000, mean reward: -0.018 [-1.000, 1.000], mean action: 2.252 [0.000, 5.000], mean observation: 105.474 [87.000, 236.000], loss: 0.010696, mean_absolute_error: 0.014047, mean_q: -0.004323, mean_eps: 0.801082
  23621/500000: episode: 25, duration: 8.749s, episode steps: 1000, steps per second: 114, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean action: 2.289 [0.000, 5.000], mean observation: 105.592 [87.000, 236.000], loss: 0.011601, mean_absolute_error: 0.015467, mean_q: -0.006702, mean_eps: 0.791902
  24497/500000: episode: 26, duration: 7.560s, episode steps: 876, steps per second: 116, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 2.507 [0.000, 5.000], mean observation: 105.441 [87.000, 236.000], loss: 0.011422, mean_absolute_error: 0.017508, mean_q: -0.009398, mean_eps: 0.783460
  25339/500000: episode: 27, duration: 7.424s, episode steps: 842, steps per second: 113, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.643 [0.000, 5.000], mean observation: 105.408 [87.000, 236.000], loss: 0.011224, mean_absolute_error: 0.019576, mean_q: -0.010343, mean_eps: 0.775738
  26191/500000: episode: 28, duration: 7.494s, episode steps: 852, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.627 [0.000, 5.000], mean observation: 105.582 [87.000, 236.000], loss: 0.009096, mean_absolute_error: 0.021657, mean_q: -0.012000, mean_eps: 0.768124
  27062/500000: episode: 29, duration: 7.654s, episode steps: 871, steps per second: 114, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.448 [0.000, 5.000], mean observation: 105.618 [87.000, 236.000], loss: 0.011286, mean_absolute_error: 0.021544, mean_q: -0.012186, mean_eps: 0.760366
  27900/500000: episode: 30, duration: 7.317s, episode steps: 838, steps per second: 115, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.912 [0.000, 5.000], mean observation: 105.423 [87.000, 236.000], loss: 0.011158, mean_absolute_error: 0.018372, mean_q: -0.008766, mean_eps: 0.752680
  28726/500000: episode: 31, duration: 7.210s, episode steps: 826, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.062 [0.000, 5.000], mean observation: 105.564 [87.000, 236.000], loss: 0.012141, mean_absolute_error: 0.021592, mean_q: -0.009837, mean_eps: 0.745192
  29634/500000: episode: 32, duration: 7.793s, episode steps: 908, steps per second: 117, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 2.137 [0.000, 5.000], mean observation: 105.420 [87.000, 236.000], loss: 0.011809, mean_absolute_error: 0.021187, mean_q: -0.012331, mean_eps: 0.737380
  30477/500000: episode: 33, duration: 7.346s, episode steps: 843, steps per second: 115, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.507 [0.000, 5.000], mean observation: 105.490 [87.000, 236.000], loss: 0.011213, mean_absolute_error: 0.019128, mean_q: -0.013170, mean_eps: 0.729496
  31241/500000: episode: 34, duration: 6.643s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.007 [0.000, 5.000], mean observation: 105.610 [87.000, 236.000], loss: 0.012676, mean_absolute_error: 0.028653, mean_q: -0.021992, mean_eps: 0.722260
  32361/500000: episode: 35, duration: 9.750s, episode steps: 1120, steps per second: 115, episode reward: -19.000, mean reward: -0.017 [-1.000, 1.000], mean action: 2.797 [0.000, 5.000], mean observation: 105.440 [87.000, 236.000], loss: 0.011901, mean_absolute_error: 0.034423, mean_q: -0.030227, mean_eps: 0.713782
  33186/500000: episode: 36, duration: 7.187s, episode steps: 825, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.701 [0.000, 5.000], mean observation: 105.577 [87.000, 236.000], loss: 0.010994, mean_absolute_error: 0.035500, mean_q: -0.030374, mean_eps: 0.705034
  34100/500000: episode: 37, duration: 8.064s, episode steps: 914, steps per second: 113, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 2.387 [0.000, 5.000], mean observation: 105.591 [87.000, 236.000], loss: 0.010747, mean_absolute_error: 0.035176, mean_q: -0.030494, mean_eps: 0.697222
  34920/500000: episode: 38, duration: 7.127s, episode steps: 820, steps per second: 115, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 2.218 [0.000, 5.000], mean observation: 105.585 [87.000, 236.000], loss: 0.012424, mean_absolute_error: 0.036220, mean_q: -0.030509, mean_eps: 0.689428
  35772/500000: episode: 39, duration: 7.382s, episode steps: 852, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.430 [0.000, 5.000], mean observation: 105.610 [87.000, 236.000], loss: 0.010877, mean_absolute_error: 0.037245, mean_q: -0.032854, mean_eps: 0.681904
  36564/500000: episode: 40, duration: 6.850s, episode steps: 792, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.352 [0.000, 5.000], mean observation: 105.584 [87.000, 236.000], loss: 0.010815, mean_absolute_error: 0.036055, mean_q: -0.032444, mean_eps: 0.674506
  37525/500000: episode: 41, duration: 8.203s, episode steps: 961, steps per second: 117, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action: 2.377 [0.000, 5.000], mean observation: 105.539 [87.000, 236.000], loss: 0.012709, mean_absolute_error: 0.037362, mean_q: -0.031812, mean_eps: 0.666604
  38351/500000: episode: 42, duration: 7.078s, episode steps: 826, steps per second: 117, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.586 [0.000, 5.000], mean observation: 105.567 [87.000, 236.000], loss: 0.010693, mean_absolute_error: 0.036643, mean_q: -0.030504, mean_eps: 0.658558
  39143/500000: episode: 43, duration: 6.783s, episode steps: 792, steps per second: 117, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.567 [0.000, 5.000], mean observation: 105.575 [87.000, 236.000], loss: 0.012718, mean_absolute_error: 0.037587, mean_q: -0.030249, mean_eps: 0.651286
  39981/500000: episode: 44, duration: 7.192s, episode steps: 838, steps per second: 117, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.879 [0.000, 5.000], mean observation: 105.604 [87.000, 236.000], loss: 0.011337, mean_absolute_error: 0.037743, mean_q: -0.032718, mean_eps: 0.643942
  40807/500000: episode: 45, duration: 7.045s, episode steps: 826, steps per second: 117, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.648 [0.000, 5.000], mean observation: 105.572 [87.000, 236.000], loss: 0.011853, mean_absolute_error: 0.041165, mean_q: -0.036931, mean_eps: 0.636454
  41571/500000: episode: 46, duration: 6.395s, episode steps: 764, steps per second: 119, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.446 [0.000, 5.000], mean observation: 105.579 [87.000, 236.000], loss: 0.011721, mean_absolute_error: 0.048252, mean_q: -0.046811, mean_eps: 0.629308
  42391/500000: episode: 47, duration: 6.898s, episode steps: 820, steps per second: 119, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 3.065 [0.000, 5.000], mean observation: 105.581 [87.000, 236.000], loss: 0.013288, mean_absolute_error: 0.054531, mean_q: -0.053949, mean_eps: 0.622180
  43353/500000: episode: 48, duration: 8.175s, episode steps: 962, steps per second: 118, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 2.983 [0.000, 5.000], mean observation: 105.454 [87.000, 236.000], loss: 0.010998, mean_absolute_error: 0.056648, mean_q: -0.055398, mean_eps: 0.614152
  44331/500000: episode: 49, duration: 8.329s, episode steps: 978, steps per second: 117, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action: 3.132 [0.000, 5.000], mean observation: 105.512 [87.000, 236.000], loss: 0.011356, mean_absolute_error: 0.055766, mean_q: -0.054141, mean_eps: 0.605422
  45123/500000: episode: 50, duration: 6.683s, episode steps: 792, steps per second: 119, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.027 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.009896, mean_absolute_error: 0.054398, mean_q: -0.051589, mean_eps: 0.597466
  45949/500000: episode: 51, duration: 6.928s, episode steps: 826, steps per second: 119, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.769 [0.000, 5.000], mean observation: 105.592 [87.000, 236.000], loss: 0.013400, mean_absolute_error: 0.055390, mean_q: -0.052579, mean_eps: 0.590176
  46792/500000: episode: 52, duration: 7.192s, episode steps: 843, steps per second: 117, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.892 [0.000, 5.000], mean observation: 105.444 [87.000, 236.000], loss: 0.012273, mean_absolute_error: 0.056403, mean_q: -0.054791, mean_eps: 0.582670
  47575/500000: episode: 53, duration: 6.734s, episode steps: 783, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.696 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.011750, mean_absolute_error: 0.054779, mean_q: -0.053962, mean_eps: 0.575362
  48457/500000: episode: 54, duration: 7.448s, episode steps: 882, steps per second: 118, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 3.235 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.012409, mean_absolute_error: 0.055952, mean_q: -0.054711, mean_eps: 0.567856
  49299/500000: episode: 55, duration: 7.245s, episode steps: 842, steps per second: 116, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 3.662 [0.000, 5.000], mean observation: 105.410 [87.000, 236.000], loss: 0.010733, mean_absolute_error: 0.057471, mean_q: -0.055524, mean_eps: 0.560098
  50137/500000: episode: 56, duration: 7.093s, episode steps: 838, steps per second: 118, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 3.578 [0.000, 5.000], mean observation: 105.386 [87.000, 236.000], loss: 0.012805, mean_absolute_error: 0.058155, mean_q: -0.056905, mean_eps: 0.552538
  50991/500000: episode: 57, duration: 7.226s, episode steps: 854, steps per second: 118, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.614 [0.000, 5.000], mean observation: 105.564 [87.000, 236.000], loss: 0.013598, mean_absolute_error: 0.068984, mean_q: -0.070039, mean_eps: 0.544924
  51938/500000: episode: 58, duration: 8.119s, episode steps: 947, steps per second: 117, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 1.992 [0.000, 5.000], mean observation: 105.473 [87.000, 236.000], loss: 0.012224, mean_absolute_error: 0.077345, mean_q: -0.078699, mean_eps: 0.536824
  52875/500000: episode: 59, duration: 8.230s, episode steps: 937, steps per second: 114, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action: 2.268 [0.000, 5.000], mean observation: 105.570 [87.000, 236.000], loss: 0.013305, mean_absolute_error: 0.079985, mean_q: -0.080165, mean_eps: 0.528346
  53639/500000: episode: 60, duration: 6.534s, episode steps: 764, steps per second: 117, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.817 [0.000, 5.000], mean observation: 105.579 [87.000, 236.000], loss: 0.012569, mean_absolute_error: 0.081340, mean_q: -0.081989, mean_eps: 0.520696
  54498/500000: episode: 61, duration: 7.378s, episode steps: 859, steps per second: 116, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 1.861 [0.000, 5.000], mean observation: 105.391 [87.000, 236.000], loss: 0.011011, mean_absolute_error: 0.078175, mean_q: -0.081169, mean_eps: 0.513388
  55322/500000: episode: 62, duration: 7.137s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.558 [0.000, 5.000], mean observation: 105.606 [87.000, 236.000], loss: 0.012814, mean_absolute_error: 0.079013, mean_q: -0.081828, mean_eps: 0.505810
  56288/500000: episode: 63, duration: 8.352s, episode steps: 966, steps per second: 116, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 2.981 [0.000, 5.000], mean observation: 105.543 [87.000, 236.000], loss: 0.012723, mean_absolute_error: 0.081053, mean_q: -0.082743, mean_eps: 0.497764
  57204/500000: episode: 64, duration: 8.033s, episode steps: 916, steps per second: 114, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 2.733 [0.000, 5.000], mean observation: 105.588 [87.000, 236.000], loss: 0.011395, mean_absolute_error: 0.077186, mean_q: -0.078418, mean_eps: 0.489304
  58030/500000: episode: 65, duration: 7.298s, episode steps: 826, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.145 [0.000, 5.000], mean observation: 105.595 [87.000, 236.000], loss: 0.013370, mean_absolute_error: 0.079945, mean_q: -0.080113, mean_eps: 0.481456
  58884/500000: episode: 66, duration: 7.476s, episode steps: 854, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.402 [0.000, 5.000], mean observation: 105.588 [87.000, 236.000], loss: 0.013469, mean_absolute_error: 0.082543, mean_q: -0.084047, mean_eps: 0.473896
  59708/500000: episode: 67, duration: 7.187s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.797 [0.000, 5.000], mean observation: 105.583 [87.000, 236.000], loss: 0.011263, mean_absolute_error: 0.080559, mean_q: -0.081033, mean_eps: 0.466354
  60612/500000: episode: 68, duration: 7.917s, episode steps: 904, steps per second: 114, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 1.850 [0.000, 5.000], mean observation: 105.547 [87.000, 236.000], loss: 0.012384, mean_absolute_error: 0.081912, mean_q: -0.086142, mean_eps: 0.458578
  61527/500000: episode: 69, duration: 8.014s, episode steps: 915, steps per second: 114, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 1.611 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.012048, mean_absolute_error: 0.091354, mean_q: -0.097280, mean_eps: 0.450388
  62351/500000: episode: 70, duration: 7.158s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.606 [0.000, 5.000], mean observation: 105.601 [87.000, 236.000], loss: 0.010103, mean_absolute_error: 0.092474, mean_q: -0.099153, mean_eps: 0.442558
  63236/500000: episode: 71, duration: 7.730s, episode steps: 885, steps per second: 114, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.468 [0.000, 5.000], mean observation: 105.617 [87.000, 236.000], loss: 0.012248, mean_absolute_error: 0.095182, mean_q: -0.100903, mean_eps: 0.434872
  64077/500000: episode: 72, duration: 7.369s, episode steps: 841, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.580 [0.000, 5.000], mean observation: 105.570 [87.000, 236.000], loss: 0.010867, mean_absolute_error: 0.095370, mean_q: -0.101026, mean_eps: 0.427096
  64920/500000: episode: 73, duration: 7.407s, episode steps: 843, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.142 [0.000, 5.000], mean observation: 105.427 [87.000, 236.000], loss: 0.012931, mean_absolute_error: 0.096512, mean_q: -0.101669, mean_eps: 0.419518
  65703/500000: episode: 74, duration: 6.881s, episode steps: 783, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.844 [0.000, 5.000], mean observation: 105.583 [87.000, 236.000], loss: 0.012328, mean_absolute_error: 0.098042, mean_q: -0.102824, mean_eps: 0.412210
  66684/500000: episode: 75, duration: 8.694s, episode steps: 981, steps per second: 113, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action: 1.620 [0.000, 5.000], mean observation: 105.531 [87.000, 236.000], loss: 0.013067, mean_absolute_error: 0.098357, mean_q: -0.104990, mean_eps: 0.404272
  67508/500000: episode: 76, duration: 7.249s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.950 [0.000, 5.000], mean observation: 105.575 [87.000, 236.000], loss: 0.011824, mean_absolute_error: 0.097908, mean_q: -0.104706, mean_eps: 0.396154
  68375/500000: episode: 77, duration: 7.695s, episode steps: 867, steps per second: 113, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 1.631 [0.000, 5.000], mean observation: 105.613 [87.000, 236.000], loss: 0.010581, mean_absolute_error: 0.095151, mean_q: -0.099810, mean_eps: 0.388540
  69199/500000: episode: 78, duration: 7.261s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.125 [0.000, 5.000], mean observation: 105.594 [87.000, 236.000], loss: 0.013613, mean_absolute_error: 0.096392, mean_q: -0.102176, mean_eps: 0.380926
  70042/500000: episode: 79, duration: 7.417s, episode steps: 843, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.422 [0.000, 5.000], mean observation: 105.574 [87.000, 236.000], loss: 0.011274, mean_absolute_error: 0.098386, mean_q: -0.105335, mean_eps: 0.373420
  70834/500000: episode: 80, duration: 6.956s, episode steps: 792, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.696 [0.000, 5.000], mean observation: 105.583 [87.000, 236.000], loss: 0.011490, mean_absolute_error: 0.102427, mean_q: -0.112077, mean_eps: 0.366058
  71598/500000: episode: 81, duration: 6.796s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.202 [0.000, 5.000], mean observation: 105.567 [87.000, 236.000], loss: 0.011933, mean_absolute_error: 0.111557, mean_q: -0.117214, mean_eps: 0.359056
  72362/500000: episode: 82, duration: 6.654s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.207 [0.000, 5.000], mean observation: 105.566 [87.000, 236.000], loss: 0.013375, mean_absolute_error: 0.118899, mean_q: -0.123015, mean_eps: 0.352180
  73186/500000: episode: 83, duration: 7.234s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.337 [0.000, 5.000], mean observation: 105.546 [87.000, 236.000], loss: 0.012660, mean_absolute_error: 0.120222, mean_q: -0.129836, mean_eps: 0.345034
  74012/500000: episode: 84, duration: 7.301s, episode steps: 826, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.971 [0.000, 5.000], mean observation: 105.554 [87.000, 236.000], loss: 0.011620, mean_absolute_error: 0.118168, mean_q: -0.127856, mean_eps: 0.337618
  74896/500000: episode: 85, duration: 7.815s, episode steps: 884, steps per second: 113, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 3.092 [0.000, 5.000], mean observation: 105.595 [87.000, 236.000], loss: 0.012081, mean_absolute_error: 0.118736, mean_q: -0.128661, mean_eps: 0.329932
  75777/500000: episode: 86, duration: 7.738s, episode steps: 881, steps per second: 114, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 1.697 [0.000, 5.000], mean observation: 105.572 [87.000, 236.000], loss: 0.011080, mean_absolute_error: 0.117478, mean_q: -0.128768, mean_eps: 0.321976
  76745/500000: episode: 87, duration: 8.464s, episode steps: 968, steps per second: 114, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 1.642 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.012595, mean_absolute_error: 0.118812, mean_q: -0.127716, mean_eps: 0.313642
  77509/500000: episode: 88, duration: 6.654s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.370 [0.000, 5.000], mean observation: 105.603 [87.000, 236.000], loss: 0.011386, mean_absolute_error: 0.114785, mean_q: -0.124609, mean_eps: 0.305848
  78273/500000: episode: 89, duration: 6.684s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.332 [0.000, 5.000], mean observation: 105.574 [87.000, 236.000], loss: 0.012078, mean_absolute_error: 0.116077, mean_q: -0.126403, mean_eps: 0.298972
  79037/500000: episode: 90, duration: 6.636s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.181 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.011695, mean_absolute_error: 0.117118, mean_q: -0.125876, mean_eps: 0.292096
  79863/500000: episode: 91, duration: 7.212s, episode steps: 826, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.861 [0.000, 5.000], mean observation: 105.545 [87.000, 236.000], loss: 0.011220, mean_absolute_error: 0.116844, mean_q: -0.128227, mean_eps: 0.284950
  80701/500000: episode: 92, duration: 7.413s, episode steps: 838, steps per second: 113, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 4.353 [0.000, 5.000], mean observation: 105.367 [87.000, 236.000], loss: 0.011167, mean_absolute_error: 0.119609, mean_q: -0.132245, mean_eps: 0.277462
  81521/500000: episode: 93, duration: 7.142s, episode steps: 820, steps per second: 115, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 1.796 [0.000, 5.000], mean observation: 105.588 [87.000, 236.000], loss: 0.014206, mean_absolute_error: 0.132704, mean_q: -0.145271, mean_eps: 0.269992
  82433/500000: episode: 94, duration: 8.017s, episode steps: 912, steps per second: 114, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 2.355 [0.000, 5.000], mean observation: 105.453 [87.000, 236.000], loss: 0.013590, mean_absolute_error: 0.138177, mean_q: -0.150581, mean_eps: 0.262198
  83197/500000: episode: 95, duration: 6.594s, episode steps: 764, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.624 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.011275, mean_absolute_error: 0.136907, mean_q: -0.150162, mean_eps: 0.254656
  83989/500000: episode: 96, duration: 7.001s, episode steps: 792, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.241 [0.000, 5.000], mean observation: 105.592 [87.000, 236.000], loss: 0.012449, mean_absolute_error: 0.139587, mean_q: -0.153578, mean_eps: 0.247654
  84815/500000: episode: 97, duration: 7.281s, episode steps: 826, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.896 [0.000, 5.000], mean observation: 105.544 [87.000, 236.000], loss: 0.010623, mean_absolute_error: 0.137366, mean_q: -0.150859, mean_eps: 0.240382
  85669/500000: episode: 98, duration: 7.500s, episode steps: 854, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.594 [0.000, 5.000], mean observation: 105.562 [87.000, 236.000], loss: 0.013547, mean_absolute_error: 0.139208, mean_q: -0.149826, mean_eps: 0.232822
  86493/500000: episode: 99, duration: 7.235s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.245 [0.000, 5.000], mean observation: 105.596 [87.000, 236.000], loss: 0.011021, mean_absolute_error: 0.138299, mean_q: -0.151580, mean_eps: 0.225262
  87276/500000: episode: 100, duration: 6.928s, episode steps: 783, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.413 [0.000, 5.000], mean observation: 105.591 [87.000, 236.000], loss: 0.013337, mean_absolute_error: 0.138109, mean_q: -0.153174, mean_eps: 0.218044
  88102/500000: episode: 101, duration: 7.326s, episode steps: 826, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.364 [0.000, 5.000], mean observation: 105.543 [87.000, 236.000], loss: 0.013607, mean_absolute_error: 0.141372, mean_q: -0.152372, mean_eps: 0.210808
  88866/500000: episode: 102, duration: 6.733s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.542 [0.000, 5.000], mean observation: 105.564 [87.000, 236.000], loss: 0.013684, mean_absolute_error: 0.144990, mean_q: -0.157989, mean_eps: 0.203644
  89690/500000: episode: 103, duration: 7.182s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.644 [0.000, 5.000], mean observation: 105.535 [87.000, 236.000], loss: 0.012468, mean_absolute_error: 0.142968, mean_q: -0.157099, mean_eps: 0.196498
  90454/500000: episode: 104, duration: 6.706s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.783 [0.000, 5.000], mean observation: 105.603 [87.000, 236.000], loss: 0.013114, mean_absolute_error: 0.143209, mean_q: -0.157283, mean_eps: 0.189352
  91374/500000: episode: 105, duration: 8.091s, episode steps: 920, steps per second: 114, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 1.274 [0.000, 5.000], mean observation: 105.418 [87.000, 236.000], loss: 0.013061, mean_absolute_error: 0.154718, mean_q: -0.171472, mean_eps: 0.181774
  92138/500000: episode: 106, duration: 6.708s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.797 [0.000, 5.000], mean observation: 105.566 [87.000, 236.000], loss: 0.012104, mean_absolute_error: 0.159669, mean_q: -0.178434, mean_eps: 0.174196
  92990/500000: episode: 107, duration: 7.540s, episode steps: 852, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.048 [0.000, 5.000], mean observation: 105.570 [87.000, 236.000], loss: 0.012729, mean_absolute_error: 0.160745, mean_q: -0.179445, mean_eps: 0.166924
  93782/500000: episode: 108, duration: 7.122s, episode steps: 792, steps per second: 111, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.255 [0.000, 5.000], mean observation: 105.569 [87.000, 236.000], loss: 0.013374, mean_absolute_error: 0.160616, mean_q: -0.179306, mean_eps: 0.159526
  94546/500000: episode: 109, duration: 6.792s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.795 [0.000, 5.000], mean observation: 105.623 [87.000, 236.000], loss: 0.014234, mean_absolute_error: 0.161916, mean_q: -0.178086, mean_eps: 0.152524
  95370/500000: episode: 110, duration: 7.243s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.081 [0.000, 5.000], mean observation: 105.626 [87.000, 236.000], loss: 0.011806, mean_absolute_error: 0.160654, mean_q: -0.177490, mean_eps: 0.145378
  96134/500000: episode: 111, duration: 6.780s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.653 [0.000, 5.000], mean observation: 105.577 [87.000, 236.000], loss: 0.012949, mean_absolute_error: 0.159968, mean_q: -0.179135, mean_eps: 0.138232
  96898/500000: episode: 112, duration: 6.718s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.424 [0.000, 5.000], mean observation: 105.602 [87.000, 236.000], loss: 0.014037, mean_absolute_error: 0.160613, mean_q: -0.177010, mean_eps: 0.131356
  97662/500000: episode: 113, duration: 6.731s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.795 [0.000, 5.000], mean observation: 105.576 [87.000, 236.000], loss: 0.014219, mean_absolute_error: 0.162720, mean_q: -0.178412, mean_eps: 0.124480
  98546/500000: episode: 114, duration: 7.827s, episode steps: 884, steps per second: 113, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 0.325 [0.000, 5.000], mean observation: 105.611 [87.000, 236.000], loss: 0.013719, mean_absolute_error: 0.162024, mean_q: -0.174992, mean_eps: 0.117064
  99449/500000: episode: 115, duration: 7.939s, episode steps: 903, steps per second: 114, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 0.282 [0.000, 5.000], mean observation: 105.608 [87.000, 236.000], loss: 0.011266, mean_absolute_error: 0.157865, mean_q: -0.172887, mean_eps: 0.109018
 100273/500000: episode: 116, duration: 7.191s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.403 [0.000, 5.000], mean observation: 105.610 [87.000, 236.000], loss: 0.011983, mean_absolute_error: 0.156101, mean_q: -0.173848, mean_eps: 0.101652
 101037/500000: episode: 117, duration: 6.726s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.024 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.013199, mean_absolute_error: 0.168712, mean_q: -0.188531, mean_eps: 0.100000
 101801/500000: episode: 118, duration: 6.727s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.797 [0.000, 5.000], mean observation: 105.579 [87.000, 236.000], loss: 0.010514, mean_absolute_error: 0.172977, mean_q: -0.192340, mean_eps: 0.100000
 102625/500000: episode: 119, duration: 7.310s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.278 [0.000, 5.000], mean observation: 105.576 [87.000, 236.000], loss: 0.012319, mean_absolute_error: 0.173778, mean_q: -0.193854, mean_eps: 0.100000
 103389/500000: episode: 120, duration: 6.662s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.588 [0.000, 5.000], mean observation: 105.594 [87.000, 236.000], loss: 0.012208, mean_absolute_error: 0.175484, mean_q: -0.196528, mean_eps: 0.100000
 104153/500000: episode: 121, duration: 6.752s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.323 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.011793, mean_absolute_error: 0.176904, mean_q: -0.197920, mean_eps: 0.100000
 104979/500000: episode: 122, duration: 7.318s, episode steps: 826, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.028 [0.000, 5.000], mean observation: 105.531 [87.000, 236.000], loss: 0.010537, mean_absolute_error: 0.175410, mean_q: -0.197143, mean_eps: 0.100000
 105803/500000: episode: 123, duration: 7.267s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.934 [0.000, 5.000], mean observation: 105.551 [87.000, 236.000], loss: 0.014517, mean_absolute_error: 0.179454, mean_q: -0.198685, mean_eps: 0.100000
 106627/500000: episode: 124, duration: 7.289s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.323 [0.000, 5.000], mean observation: 105.558 [87.000, 236.000], loss: 0.013489, mean_absolute_error: 0.181773, mean_q: -0.202659, mean_eps: 0.100000
 107451/500000: episode: 125, duration: 7.278s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.706 [0.000, 5.000], mean observation: 105.586 [87.000, 236.000], loss: 0.013075, mean_absolute_error: 0.178587, mean_q: -0.200423, mean_eps: 0.100000
 108275/500000: episode: 126, duration: 7.281s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.021 [0.000, 5.000], mean observation: 105.580 [87.000, 236.000], loss: 0.012213, mean_absolute_error: 0.181184, mean_q: -0.199854, mean_eps: 0.100000
 109161/500000: episode: 127, duration: 7.824s, episode steps: 886, steps per second: 113, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 0.754 [0.000, 5.000], mean observation: 105.586 [87.000, 236.000], loss: 0.011650, mean_absolute_error: 0.176577, mean_q: -0.195379, mean_eps: 0.100000
 110045/500000: episode: 128, duration: 7.791s, episode steps: 884, steps per second: 113, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 1.428 [0.000, 5.000], mean observation: 105.608 [87.000, 236.000], loss: 0.011683, mean_absolute_error: 0.175434, mean_q: -0.198023, mean_eps: 0.100000
 110869/500000: episode: 129, duration: 7.259s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.499 [0.000, 5.000], mean observation: 105.554 [87.000, 236.000], loss: 0.012525, mean_absolute_error: 0.183101, mean_q: -0.208038, mean_eps: 0.100000
 111695/500000: episode: 130, duration: 7.317s, episode steps: 826, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.908 [0.000, 5.000], mean observation: 105.553 [87.000, 236.000], loss: 0.012272, mean_absolute_error: 0.193673, mean_q: -0.219695, mean_eps: 0.100000
 112519/500000: episode: 131, duration: 7.261s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.041 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.011594, mean_absolute_error: 0.193381, mean_q: -0.219836, mean_eps: 0.100000
 113311/500000: episode: 132, duration: 6.997s, episode steps: 792, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.543 [0.000, 5.000], mean observation: 105.567 [87.000, 236.000], loss: 0.011621, mean_absolute_error: 0.192219, mean_q: -0.217508, mean_eps: 0.100000
 114075/500000: episode: 133, duration: 6.650s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.949 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.011670, mean_absolute_error: 0.193114, mean_q: -0.216259, mean_eps: 0.100000
 115114/500000: episode: 134, duration: 9.106s, episode steps: 1039, steps per second: 114, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action: 1.407 [0.000, 5.000], mean observation: 105.484 [87.000, 236.000], loss: 0.013238, mean_absolute_error: 0.196815, mean_q: -0.219639, mean_eps: 0.100000
 115938/500000: episode: 135, duration: 7.231s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.894 [0.000, 5.000], mean observation: 105.561 [87.000, 236.000], loss: 0.010585, mean_absolute_error: 0.195933, mean_q: -0.218388, mean_eps: 0.100000
 116702/500000: episode: 136, duration: 6.765s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.177 [0.000, 5.000], mean observation: 105.564 [87.000, 236.000], loss: 0.010866, mean_absolute_error: 0.194185, mean_q: -0.215857, mean_eps: 0.100000
 117528/500000: episode: 137, duration: 7.203s, episode steps: 826, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.435 [0.000, 5.000], mean observation: 105.551 [87.000, 236.000], loss: 0.013391, mean_absolute_error: 0.196973, mean_q: -0.218931, mean_eps: 0.100000
 118292/500000: episode: 138, duration: 6.765s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.055 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.012619, mean_absolute_error: 0.197051, mean_q: -0.219893, mean_eps: 0.100000
 119116/500000: episode: 139, duration: 7.257s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.938 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.012838, mean_absolute_error: 0.197062, mean_q: -0.220641, mean_eps: 0.100000
 119880/500000: episode: 140, duration: 6.692s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.368 [0.000, 5.000], mean observation: 105.564 [87.000, 236.000], loss: 0.012553, mean_absolute_error: 0.197000, mean_q: -0.219471, mean_eps: 0.100000
 120644/500000: episode: 141, duration: 6.731s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.520 [0.000, 5.000], mean observation: 105.569 [87.000, 236.000], loss: 0.012152, mean_absolute_error: 0.200263, mean_q: -0.230467, mean_eps: 0.100000
 121436/500000: episode: 142, duration: 7.010s, episode steps: 792, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.996 [0.000, 5.000], mean observation: 105.552 [87.000, 236.000], loss: 0.010741, mean_absolute_error: 0.210830, mean_q: -0.242102, mean_eps: 0.100000
 122200/500000: episode: 143, duration: 6.886s, episode steps: 764, steps per second: 111, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.949 [0.000, 5.000], mean observation: 105.578 [87.000, 236.000], loss: 0.012333, mean_absolute_error: 0.214021, mean_q: -0.242806, mean_eps: 0.100000
 123054/500000: episode: 144, duration: 7.525s, episode steps: 854, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.680 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.013836, mean_absolute_error: 0.219526, mean_q: -0.248354, mean_eps: 0.100000
 123874/500000: episode: 145, duration: 7.154s, episode steps: 820, steps per second: 115, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 1.587 [0.000, 5.000], mean observation: 105.600 [87.000, 236.000], loss: 0.011284, mean_absolute_error: 0.218048, mean_q: -0.244349, mean_eps: 0.100000
 124716/500000: episode: 146, duration: 7.368s, episode steps: 842, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.198 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.011716, mean_absolute_error: 0.214533, mean_q: -0.243576, mean_eps: 0.100000
 125635/500000: episode: 147, duration: 8.137s, episode steps: 919, steps per second: 113, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 1.618 [0.000, 5.000], mean observation: 105.571 [87.000, 236.000], loss: 0.012368, mean_absolute_error: 0.217731, mean_q: -0.246191, mean_eps: 0.100000
 126399/500000: episode: 148, duration: 6.699s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.147 [0.000, 5.000], mean observation: 105.598 [87.000, 236.000], loss: 0.012440, mean_absolute_error: 0.217500, mean_q: -0.245651, mean_eps: 0.100000
 127163/500000: episode: 149, duration: 6.724s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.132 [0.000, 5.000], mean observation: 105.570 [87.000, 236.000], loss: 0.014555, mean_absolute_error: 0.222612, mean_q: -0.249300, mean_eps: 0.100000
 127927/500000: episode: 150, duration: 6.892s, episode steps: 764, steps per second: 111, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.255 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.012398, mean_absolute_error: 0.217661, mean_q: -0.243681, mean_eps: 0.100000
 128691/500000: episode: 151, duration: 7.268s, episode steps: 764, steps per second: 105, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.613 [0.000, 5.000], mean observation: 105.600 [87.000, 236.000], loss: 0.011995, mean_absolute_error: 0.215281, mean_q: -0.243014, mean_eps: 0.100000
 129455/500000: episode: 152, duration: 6.743s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.822 [0.000, 5.000], mean observation: 105.578 [87.000, 236.000], loss: 0.010583, mean_absolute_error: 0.213570, mean_q: -0.243350, mean_eps: 0.100000
 130219/500000: episode: 153, duration: 6.682s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.507 [0.000, 5.000], mean observation: 105.594 [87.000, 236.000], loss: 0.012961, mean_absolute_error: 0.215679, mean_q: -0.246025, mean_eps: 0.100000
 131225/500000: episode: 154, duration: 8.807s, episode steps: 1006, steps per second: 114, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean action: 1.018 [0.000, 5.000], mean observation: 105.614 [87.000, 236.000], loss: 0.013573, mean_absolute_error: 0.229974, mean_q: -0.262066, mean_eps: 0.100000
 132017/500000: episode: 155, duration: 7.058s, episode steps: 792, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.529 [0.000, 5.000], mean observation: 105.612 [87.000, 236.000], loss: 0.011479, mean_absolute_error: 0.232517, mean_q: -0.265697, mean_eps: 0.100000
 132999/500000: episode: 156, duration: 8.715s, episode steps: 982, steps per second: 113, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action: 0.901 [0.000, 5.000], mean observation: 105.611 [87.000, 236.000], loss: 0.012646, mean_absolute_error: 0.234668, mean_q: -0.266813, mean_eps: 0.100000
 133859/500000: episode: 157, duration: 7.575s, episode steps: 860, steps per second: 114, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 0.270 [0.000, 5.000], mean observation: 105.421 [87.000, 236.000], loss: 0.013433, mean_absolute_error: 0.235884, mean_q: -0.262732, mean_eps: 0.100000
 134822/500000: episode: 158, duration: 8.451s, episode steps: 963, steps per second: 114, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 0.271 [0.000, 5.000], mean observation: 105.505 [87.000, 236.000], loss: 0.011768, mean_absolute_error: 0.234880, mean_q: -0.262062, mean_eps: 0.100000
 135646/500000: episode: 159, duration: 7.253s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.567 [0.000, 5.000], mean observation: 105.600 [87.000, 236.000], loss: 0.012573, mean_absolute_error: 0.234937, mean_q: -0.264657, mean_eps: 0.100000
 136466/500000: episode: 160, duration: 7.204s, episode steps: 820, steps per second: 114, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 1.420 [0.000, 5.000], mean observation: 105.589 [87.000, 236.000], loss: 0.013177, mean_absolute_error: 0.235325, mean_q: -0.267215, mean_eps: 0.100000
 137442/500000: episode: 161, duration: 8.515s, episode steps: 976, steps per second: 115, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action: 0.864 [0.000, 5.000], mean observation: 105.442 [87.000, 236.000], loss: 0.011434, mean_absolute_error: 0.234579, mean_q: -0.264541, mean_eps: 0.100000
 138266/500000: episode: 162, duration: 7.243s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.379 [0.000, 5.000], mean observation: 105.619 [87.000, 236.000], loss: 0.011035, mean_absolute_error: 0.231812, mean_q: -0.262408, mean_eps: 0.100000
 139092/500000: episode: 163, duration: 7.224s, episode steps: 826, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.987 [0.000, 5.000], mean observation: 105.597 [87.000, 236.000], loss: 0.010941, mean_absolute_error: 0.231761, mean_q: -0.263207, mean_eps: 0.100000
 139916/500000: episode: 164, duration: 7.263s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.201 [0.000, 5.000], mean observation: 105.601 [87.000, 236.000], loss: 0.011145, mean_absolute_error: 0.230993, mean_q: -0.263083, mean_eps: 0.100000
 140708/500000: episode: 165, duration: 6.913s, episode steps: 792, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.712 [0.000, 5.000], mean observation: 105.548 [87.000, 236.000], loss: 0.011854, mean_absolute_error: 0.236625, mean_q: -0.272521, mean_eps: 0.100000
 141532/500000: episode: 166, duration: 7.196s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.393 [0.000, 5.000], mean observation: 105.577 [87.000, 236.000], loss: 0.011973, mean_absolute_error: 0.247797, mean_q: -0.283543, mean_eps: 0.100000
 142712/500000: episode: 167, duration: 10.430s, episode steps: 1180, steps per second: 113, episode reward: -18.000, mean reward: -0.015 [-1.000, 1.000], mean action: 1.609 [0.000, 5.000], mean observation: 105.530 [87.000, 236.000], loss: 0.011569, mean_absolute_error: 0.249176, mean_q: -0.285325, mean_eps: 0.100000
 143776/500000: episode: 168, duration: 9.226s, episode steps: 1064, steps per second: 115, episode reward: -21.000, mean reward: -0.020 [-1.000, 0.000], mean action: 0.221 [0.000, 5.000], mean observation: 105.611 [87.000, 236.000], loss: 0.010839, mean_absolute_error: 0.247600, mean_q: -0.279858, mean_eps: 0.100000
 144722/500000: episode: 169, duration: 8.262s, episode steps: 946, steps per second: 114, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 0.449 [0.000, 5.000], mean observation: 105.601 [87.000, 236.000], loss: 0.011293, mean_absolute_error: 0.249462, mean_q: -0.282308, mean_eps: 0.100000
 145546/500000: episode: 170, duration: 7.275s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.820 [0.000, 5.000], mean observation: 105.611 [87.000, 236.000], loss: 0.011670, mean_absolute_error: 0.248706, mean_q: -0.280214, mean_eps: 0.100000
 146372/500000: episode: 171, duration: 7.302s, episode steps: 826, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.673 [0.000, 5.000], mean observation: 105.579 [87.000, 236.000], loss: 0.009684, mean_absolute_error: 0.248567, mean_q: -0.283780, mean_eps: 0.100000
 147344/500000: episode: 172, duration: 8.542s, episode steps: 972, steps per second: 114, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 1.238 [0.000, 5.000], mean observation: 105.607 [87.000, 236.000], loss: 0.012907, mean_absolute_error: 0.254086, mean_q: -0.290239, mean_eps: 0.100000
 148108/500000: episode: 173, duration: 6.685s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.264 [0.000, 5.000], mean observation: 105.629 [87.000, 236.000], loss: 0.011935, mean_absolute_error: 0.252142, mean_q: -0.285818, mean_eps: 0.100000
 148969/500000: episode: 174, duration: 7.528s, episode steps: 861, steps per second: 114, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 2.928 [0.000, 5.000], mean observation: 105.385 [87.000, 236.000], loss: 0.011657, mean_absolute_error: 0.250709, mean_q: -0.286206, mean_eps: 0.100000
 149793/500000: episode: 175, duration: 7.600s, episode steps: 824, steps per second: 108, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.360 [0.000, 5.000], mean observation: 105.627 [87.000, 236.000], loss: 0.011692, mean_absolute_error: 0.252722, mean_q: -0.286899, mean_eps: 0.100000
 150557/500000: episode: 176, duration: 6.715s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.681 [0.000, 5.000], mean observation: 105.633 [87.000, 236.000], loss: 0.010446, mean_absolute_error: 0.251978, mean_q: -0.291053, mean_eps: 0.100000
 151321/500000: episode: 177, duration: 6.955s, episode steps: 764, steps per second: 110, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.771 [0.000, 5.000], mean observation: 105.572 [87.000, 236.000], loss: 0.010936, mean_absolute_error: 0.262943, mean_q: -0.303494, mean_eps: 0.100000
 152085/500000: episode: 178, duration: 6.693s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.861 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.010959, mean_absolute_error: 0.263257, mean_q: -0.302906, mean_eps: 0.100000
 152849/500000: episode: 179, duration: 6.738s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.950 [0.000, 5.000], mean observation: 105.597 [87.000, 236.000], loss: 0.012158, mean_absolute_error: 0.268120, mean_q: -0.305624, mean_eps: 0.100000
 153673/500000: episode: 180, duration: 7.270s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.754 [0.000, 5.000], mean observation: 105.584 [87.000, 236.000], loss: 0.011475, mean_absolute_error: 0.267262, mean_q: -0.305263, mean_eps: 0.100000
 154639/500000: episode: 181, duration: 8.462s, episode steps: 966, steps per second: 114, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action: 1.391 [0.000, 5.000], mean observation: 105.425 [87.000, 236.000], loss: 0.009920, mean_absolute_error: 0.263760, mean_q: -0.304069, mean_eps: 0.100000
 155403/500000: episode: 182, duration: 6.790s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.474 [0.000, 5.000], mean observation: 105.600 [87.000, 236.000], loss: 0.011137, mean_absolute_error: 0.263300, mean_q: -0.302695, mean_eps: 0.100000
 156167/500000: episode: 183, duration: 6.927s, episode steps: 764, steps per second: 110, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.062 [0.000, 5.000], mean observation: 105.564 [87.000, 236.000], loss: 0.010702, mean_absolute_error: 0.265633, mean_q: -0.305286, mean_eps: 0.100000
 156979/500000: episode: 184, duration: 7.485s, episode steps: 812, steps per second: 108, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 2.070 [0.000, 5.000], mean observation: 105.601 [87.000, 236.000], loss: 0.012200, mean_absolute_error: 0.268194, mean_q: -0.304899, mean_eps: 0.100000
 157743/500000: episode: 185, duration: 6.893s, episode steps: 764, steps per second: 111, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.520 [0.000, 5.000], mean observation: 105.614 [87.000, 236.000], loss: 0.011386, mean_absolute_error: 0.267530, mean_q: -0.306986, mean_eps: 0.100000
 158537/500000: episode: 186, duration: 7.025s, episode steps: 794, steps per second: 113, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 2.664 [0.000, 5.000], mean observation: 105.550 [87.000, 236.000], loss: 0.011885, mean_absolute_error: 0.270140, mean_q: -0.309570, mean_eps: 0.100000
 159301/500000: episode: 187, duration: 6.824s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.720 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.012569, mean_absolute_error: 0.269006, mean_q: -0.305211, mean_eps: 0.100000
 160065/500000: episode: 188, duration: 6.750s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.847 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.011579, mean_absolute_error: 0.268946, mean_q: -0.303686, mean_eps: 0.100000
 160829/500000: episode: 189, duration: 6.783s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.257 [0.000, 5.000], mean observation: 105.569 [87.000, 236.000], loss: 0.011397, mean_absolute_error: 0.272302, mean_q: -0.315160, mean_eps: 0.100000
 161593/500000: episode: 190, duration: 6.781s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.288 [0.000, 5.000], mean observation: 105.612 [87.000, 236.000], loss: 0.013637, mean_absolute_error: 0.286402, mean_q: -0.329351, mean_eps: 0.100000
 162357/500000: episode: 191, duration: 6.865s, episode steps: 764, steps per second: 111, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.359 [0.000, 5.000], mean observation: 105.589 [87.000, 236.000], loss: 0.012154, mean_absolute_error: 0.286751, mean_q: -0.329393, mean_eps: 0.100000
 163301/500000: episode: 192, duration: 8.649s, episode steps: 944, steps per second: 109, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 1.799 [0.000, 5.000], mean observation: 105.577 [87.000, 236.000], loss: 0.013587, mean_absolute_error: 0.287435, mean_q: -0.328340, mean_eps: 0.100000
 164127/500000: episode: 193, duration: 7.535s, episode steps: 826, steps per second: 110, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.243 [0.000, 5.000], mean observation: 105.556 [87.000, 236.000], loss: 0.012325, mean_absolute_error: 0.288402, mean_q: -0.331517, mean_eps: 0.100000
 164891/500000: episode: 194, duration: 6.788s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.760 [0.000, 5.000], mean observation: 105.608 [87.000, 236.000], loss: 0.013014, mean_absolute_error: 0.288453, mean_q: -0.331007, mean_eps: 0.100000
 165655/500000: episode: 195, duration: 6.819s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.546 [0.000, 5.000], mean observation: 105.576 [87.000, 236.000], loss: 0.013651, mean_absolute_error: 0.289737, mean_q: -0.331395, mean_eps: 0.100000
 166557/500000: episode: 196, duration: 8.037s, episode steps: 902, steps per second: 112, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 2.060 [0.000, 5.000], mean observation: 105.407 [87.000, 236.000], loss: 0.011632, mean_absolute_error: 0.284449, mean_q: -0.326377, mean_eps: 0.100000
 167349/500000: episode: 197, duration: 7.036s, episode steps: 792, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.501 [0.000, 5.000], mean observation: 105.573 [87.000, 236.000], loss: 0.011148, mean_absolute_error: 0.278105, mean_q: -0.319924, mean_eps: 0.100000
 168210/500000: episode: 198, duration: 7.703s, episode steps: 861, steps per second: 112, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 0.317 [0.000, 5.000], mean observation: 105.433 [87.000, 236.000], loss: 0.013118, mean_absolute_error: 0.285216, mean_q: -0.325372, mean_eps: 0.100000
 169034/500000: episode: 199, duration: 7.360s, episode steps: 824, steps per second: 112, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.096 [0.000, 5.000], mean observation: 105.564 [87.000, 236.000], loss: 0.014446, mean_absolute_error: 0.291268, mean_q: -0.332599, mean_eps: 0.100000
 169978/500000: episode: 200, duration: 8.397s, episode steps: 944, steps per second: 112, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 0.321 [0.000, 5.000], mean observation: 105.589 [87.000, 236.000], loss: 0.012226, mean_absolute_error: 0.289087, mean_q: -0.329380, mean_eps: 0.100000
 170771/500000: episode: 201, duration: 7.068s, episode steps: 793, steps per second: 112, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 1.994 [0.000, 5.000], mean observation: 105.608 [87.000, 236.000], loss: 0.011740, mean_absolute_error: 0.291707, mean_q: -0.338110, mean_eps: 0.100000
 171535/500000: episode: 202, duration: 6.766s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.645 [0.000, 5.000], mean observation: 105.599 [87.000, 236.000], loss: 0.011927, mean_absolute_error: 0.300314, mean_q: -0.346538, mean_eps: 0.100000
 172327/500000: episode: 203, duration: 7.061s, episode steps: 792, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.158 [0.000, 5.000], mean observation: 105.573 [87.000, 236.000], loss: 0.012991, mean_absolute_error: 0.307381, mean_q: -0.353335, mean_eps: 0.100000
 173165/500000: episode: 204, duration: 7.464s, episode steps: 838, steps per second: 112, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.696 [0.000, 5.000], mean observation: 105.377 [87.000, 236.000], loss: 0.012935, mean_absolute_error: 0.306058, mean_q: -0.350465, mean_eps: 0.100000
 173929/500000: episode: 205, duration: 6.727s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.478 [0.000, 5.000], mean observation: 105.595 [87.000, 236.000], loss: 0.012897, mean_absolute_error: 0.306450, mean_q: -0.349928, mean_eps: 0.100000
 174693/500000: episode: 206, duration: 6.821s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.534 [0.000, 5.000], mean observation: 105.599 [87.000, 236.000], loss: 0.012363, mean_absolute_error: 0.306211, mean_q: -0.351287, mean_eps: 0.100000
 175614/500000: episode: 207, duration: 8.207s, episode steps: 921, steps per second: 112, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 1.790 [0.000, 5.000], mean observation: 105.542 [87.000, 236.000], loss: 0.012276, mean_absolute_error: 0.303491, mean_q: -0.350144, mean_eps: 0.100000
 176378/500000: episode: 208, duration: 6.863s, episode steps: 764, steps per second: 111, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.378 [0.000, 5.000], mean observation: 105.604 [87.000, 236.000], loss: 0.010859, mean_absolute_error: 0.302328, mean_q: -0.349590, mean_eps: 0.100000
 177142/500000: episode: 209, duration: 6.805s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.329 [0.000, 5.000], mean observation: 105.602 [87.000, 236.000], loss: 0.011677, mean_absolute_error: 0.301615, mean_q: -0.346983, mean_eps: 0.100000
 177994/500000: episode: 210, duration: 7.544s, episode steps: 852, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.269 [0.000, 5.000], mean observation: 105.553 [87.000, 236.000], loss: 0.013547, mean_absolute_error: 0.305061, mean_q: -0.349854, mean_eps: 0.100000
 178758/500000: episode: 211, duration: 6.736s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.004 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.012837, mean_absolute_error: 0.307827, mean_q: -0.348652, mean_eps: 0.100000
 179522/500000: episode: 212, duration: 6.818s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.809 [0.000, 5.000], mean observation: 105.578 [87.000, 236.000], loss: 0.011664, mean_absolute_error: 0.306626, mean_q: -0.351547, mean_eps: 0.100000
 180443/500000: episode: 213, duration: 8.144s, episode steps: 921, steps per second: 113, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 0.502 [0.000, 5.000], mean observation: 105.438 [87.000, 236.000], loss: 0.011509, mean_absolute_error: 0.304749, mean_q: -0.350694, mean_eps: 0.100000
```

```
181269/500000: episode: 214, duration: 7.302s, episode steps: 826, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.594 [0.000, 5.000], mean observation: 105.566 [87.000, 236.000], loss: 0.013029, mean_absolute_error: 0.319832, mean_q: -0.368936, mean_eps: 0.100000
182095/500000: episode: 215, duration: 7.263s, episode steps: 826, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.253 [0.000, 5.000], mean observation: 105.613 [87.000, 236.000], loss: 0.013310, mean_absolute_error: 0.326555, mean_q: -0.374104, mean_eps: 0.100000
182859/500000: episode: 216, duration: 6.795s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.652 [0.000, 5.000], mean observation: 105.586 [87.000, 236.000], loss: 0.011393, mean_absolute_error: 0.319774, mean_q: -0.369002, mean_eps: 0.100000
183623/500000: episode: 217, duration: 6.762s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.588 [0.000, 5.000], mean observation: 105.566 [87.000, 236.000], loss: 0.013457, mean_absolute_error: 0.321872, mean_q: -0.371044, mean_eps: 0.100000
184387/500000: episode: 218, duration: 6.738s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.018 [0.000, 5.000], mean observation: 105.580 [87.000, 236.000], loss: 0.013011, mean_absolute_error: 0.323646, mean_q: -0.372820, mean_eps: 0.100000
185231/500000: episode: 219, duration: 7.527s, episode steps: 844, steps per second: 112, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.014 [0.000, 5.000], mean observation: 105.492 [87.000, 236.000], loss: 0.011492, mean_absolute_error: 0.322269, mean_q: -0.372357, mean_eps: 0.100000
186175/500000: episode: 220, duration: 8.433s, episode steps: 944, steps per second: 112, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 0.528 [0.000, 5.000], mean observation: 105.666 [87.000, 236.000], loss: 0.012244, mean_absolute_error: 0.319353, mean_q: -0.365232, mean_eps: 0.100000
186967/500000: episode: 221, duration: 7.040s, episode steps: 792, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.218 [0.000, 5.000], mean observation: 105.590 [87.000, 236.000], loss: 0.011753, mean_absolute_error: 0.317578, mean_q: -0.366308, mean_eps: 0.100000
187791/500000: episode: 222, duration: 7.290s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.471 [0.000, 5.000], mean observation: 105.580 [87.000, 236.000], loss: 0.011051, mean_absolute_error: 0.320803, mean_q: -0.369158, mean_eps: 0.100000
188555/500000: episode: 223, duration: 6.770s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.208 [0.000, 5.000], mean observation: 105.623 [87.000, 236.000], loss: 0.011997, mean_absolute_error: 0.319234, mean_q: -0.365842, mean_eps: 0.100000
189379/500000: episode: 224, duration: 7.313s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.040 [0.000, 5.000], mean observation: 105.551 [87.000, 236.000], loss: 0.013155, mean_absolute_error: 0.319813, mean_q: -0.367191, mean_eps: 0.100000
190143/500000: episode: 225, duration: 6.781s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.495 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.012309, mean_absolute_error: 0.322391, mean_q: -0.372411, mean_eps: 0.100000
190907/500000: episode: 226, duration: 7.017s, episode steps: 764, steps per second: 109, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 4.675 [0.000, 5.000], mean observation: 105.569 [87.000, 236.000], loss: 0.012982, mean_absolute_error: 0.332666, mean_q: -0.383071, mean_eps: 0.100000
191690/500000: episode: 227, duration: 6.977s, episode steps: 783, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 4.083 [0.000, 5.000], mean observation: 105.544 [87.000, 236.000], loss: 0.014037, mean_absolute_error: 0.339180, mean_q: -0.386451, mean_eps: 0.100000
192792/500000: episode: 228, duration: 9.691s, episode steps: 1102, steps per second: 114, episode reward: -19.000, mean reward: -0.017 [-1.000, 1.000], mean action: 1.450 [0.000, 5.000], mean observation: 105.552 [87.000, 236.000], loss: 0.012776, mean_absolute_error: 0.340277, mean_q: -0.391091, mean_eps: 0.100000
193764/500000: episode: 229, duration: 8.571s, episode steps: 972, steps per second: 113, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 0.557 [0.000, 5.000], mean observation: 105.615 [87.000, 236.000], loss: 0.011764, mean_absolute_error: 0.337552, mean_q: -0.388142, mean_eps: 0.100000
194588/500000: episode: 230, duration: 7.276s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.458 [0.000, 5.000], mean observation: 105.622 [87.000, 236.000], loss: 0.011640, mean_absolute_error: 0.335395, mean_q: -0.385824, mean_eps: 0.100000
195352/500000: episode: 231, duration: 6.822s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.800 [0.000, 5.000], mean observation: 105.624 [87.000, 236.000], loss: 0.011243, mean_absolute_error: 0.335474, mean_q: -0.386922, mean_eps: 0.100000
196116/500000: episode: 232, duration: 6.833s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.685 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.012391, mean_absolute_error: 0.340988, mean_q: -0.391032, mean_eps: 0.100000
196880/500000: episode: 233, duration: 6.808s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.372 [0.000, 5.000], mean observation: 105.578 [87.000, 236.000], loss: 0.012964, mean_absolute_error: 0.343109, mean_q: -0.396267, mean_eps: 0.100000
197644/500000: episode: 234, duration: 6.890s, episode steps: 764, steps per second: 111, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.009 [0.000, 5.000], mean observation: 105.569 [87.000, 236.000], loss: 0.012643, mean_absolute_error: 0.339262, mean_q: -0.393396, mean_eps: 0.100000
198528/500000: episode: 235, duration: 8.182s, episode steps: 884, steps per second: 108, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 0.525 [0.000, 5.000], mean observation: 105.573 [87.000, 236.000], loss: 0.009556, mean_absolute_error: 0.335311, mean_q: -0.389371, mean_eps: 0.100000
199292/500000: episode: 236, duration: 6.686s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.639 [0.000, 5.000], mean observation: 105.613 [87.000, 236.000], loss: 0.009812, mean_absolute_error: 0.331546, mean_q: -0.385228, mean_eps: 0.100000
200116/500000: episode: 237, duration: 7.306s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.604 [0.000, 5.000], mean observation: 105.603 [87.000, 236.000], loss: 0.011406, mean_absolute_error: 0.338408, mean_q: -0.392740, mean_eps: 0.100000
200940/500000: episode: 238, duration: 7.293s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.697 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.011731, mean_absolute_error: 0.344556, mean_q: -0.400099, mean_eps: 0.100000
201704/500000: episode: 239, duration: 6.749s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.351 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.013234, mean_absolute_error: 0.352413, mean_q: -0.407118, mean_eps: 0.100000
202468/500000: episode: 240, duration: 6.743s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.783 [0.000, 5.000], mean observation: 105.564 [87.000, 236.000], loss: 0.014521, mean_absolute_error: 0.358331, mean_q: -0.409760, mean_eps: 0.100000
203292/500000: episode: 241, duration: 7.319s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.239 [0.000, 5.000], mean observation: 105.625 [87.000, 236.000], loss: 0.010607, mean_absolute_error: 0.351346, mean_q: -0.400637, mean_eps: 0.100000
204056/500000: episode: 242, duration: 6.705s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.609 [0.000, 5.000], mean observation: 105.606 [87.000, 236.000], loss: 0.012673, mean_absolute_error: 0.352061, mean_q: -0.409597, mean_eps: 0.100000
204820/500000: episode: 243, duration: 6.736s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.683 [0.000, 5.000], mean observation: 105.600 [87.000, 236.000], loss: 0.009464, mean_absolute_error: 0.348728, mean_q: -0.403470, mean_eps: 0.100000
205584/500000: episode: 244, duration: 6.770s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.809 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.012765, mean_absolute_error: 0.350870, mean_q: -0.404069, mean_eps: 0.100000
206468/500000: episode: 245, duration: 7.869s, episode steps: 884, steps per second: 112, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 1.605 [0.000, 5.000], mean observation: 105.596 [87.000, 236.000], loss: 0.011330, mean_absolute_error: 0.352065, mean_q: -0.405376, mean_eps: 0.100000
207232/500000: episode: 246, duration: 6.772s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.297 [0.000, 5.000], mean observation: 105.620 [87.000, 236.000], loss: 0.011453, mean_absolute_error: 0.349188, mean_q: -0.403766, mean_eps: 0.100000
208176/500000: episode: 247, duration: 8.359s, episode steps: 944, steps per second: 113, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 1.760 [0.000, 5.000], mean observation: 105.572 [87.000, 236.000], loss: 0.013006, mean_absolute_error: 0.353539, mean_q: -0.409693, mean_eps: 0.100000
209000/500000: episode: 248, duration: 7.404s, episode steps: 824, steps per second: 111, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.004 [0.000, 5.000], mean observation: 105.582 [87.000, 236.000], loss: 0.012603, mean_absolute_error: 0.352661, mean_q: -0.407828, mean_eps: 0.100000
209838/500000: episode: 249, duration: 7.482s, episode steps: 838, steps per second: 112, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 3.118 [0.000, 5.000], mean observation: 105.388 [87.000, 236.000], loss: 0.011373, mean_absolute_error: 0.353829, mean_q: -0.408394, mean_eps: 0.100000
210630/500000: episode: 250, duration: 7.007s, episode steps: 792, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.525 [0.000, 5.000], mean observation: 105.601 [87.000, 236.000], loss: 0.011732, mean_absolute_error: 0.355333, mean_q: -0.412837, mean_eps: 0.100000
211472/500000: episode: 251, duration: 7.352s, episode steps: 842, steps per second: 115, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 4.761 [0.000, 5.000], mean observation: 105.408 [87.000, 236.000], loss: 0.013056, mean_absolute_error: 0.366746, mean_q: -0.423893, mean_eps: 0.100000
212264/500000: episode: 252, duration: 6.967s, episode steps: 792, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.468 [0.000, 5.000], mean observation: 105.559 [87.000, 236.000], loss: 0.012422, mean_absolute_error: 0.369124, mean_q: -0.428385, mean_eps: 0.100000
213056/500000: episode: 253, duration: 6.982s, episode steps: 792, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.246 [0.000, 5.000], mean observation: 105.593 [87.000, 236.000], loss: 0.010733, mean_absolute_error: 0.365034, mean_q: -0.423162, mean_eps: 0.100000
213820/500000: episode: 254, duration: 6.801s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.919 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.012227, mean_absolute_error: 0.368021, mean_q: -0.427235, mean_eps: 0.100000
214603/500000: episode: 255, duration: 6.892s, episode steps: 783, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.443 [0.000, 5.000], mean observation: 105.561 [87.000, 236.000], loss: 0.013002, mean_absolute_error: 0.371882, mean_q: -0.429661, mean_eps: 0.100000
215367/500000: episode: 256, duration: 6.731s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 4.785 [0.000, 5.000], mean observation: 105.567 [87.000, 236.000], loss: 0.012226, mean_absolute_error: 0.370596, mean_q: -0.424871, mean_eps: 0.100000
216131/500000: episode: 257, duration: 6.740s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.797 [0.000, 5.000], mean observation: 105.575 [87.000, 236.000], loss: 0.013436, mean_absolute_error: 0.372042, mean_q: -0.428189, mean_eps: 0.100000
217029/500000: episode: 258, duration: 7.937s, episode steps: 898, steps per second: 113, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 4.298 [0.000, 5.000], mean observation: 105.469 [87.000, 236.000], loss: 0.012775, mean_absolute_error: 0.372400, mean_q: -0.426506, mean_eps: 0.100000
217853/500000: episode: 259, duration: 7.245s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.309 [0.000, 5.000], mean observation: 105.581 [87.000, 236.000], loss: 0.011230, mean_absolute_error: 0.370389, mean_q: -0.426189, mean_eps: 0.100000
218695/500000: episode: 260, duration: 7.272s, episode steps: 842, steps per second: 116, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 4.701 [0.000, 5.000], mean observation: 105.464 [87.000, 236.000], loss: 0.011289, mean_absolute_error: 0.368522, mean_q: -0.425080, mean_eps: 0.100000
219519/500000: episode: 261, duration: 7.223s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.766 [0.000, 5.000], mean observation: 105.582 [87.000, 236.000], loss: 0.013057, mean_absolute_error: 0.369524, mean_q: -0.426824, mean_eps: 0.100000
220343/500000: episode: 262, duration: 7.192s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.428 [0.000, 5.000], mean observation: 105.555 [87.000, 236.000], loss: 0.011392, mean_absolute_error: 0.368856, mean_q: -0.426078, mean_eps: 0.100000
221107/500000: episode: 263, duration: 6.739s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.046 [0.000, 5.000], mean observation: 105.569 [87.000, 236.000], loss: 0.011915, mean_absolute_error: 0.379623, mean_q: -0.441120, mean_eps: 0.100000
221931/500000: episode: 264, duration: 7.217s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 4.068 [0.000, 5.000], mean observation: 105.551 [87.000, 236.000], loss: 0.011362, mean_absolute_error: 0.382247, mean_q: -0.441602, mean_eps: 0.100000
222695/500000: episode: 265, duration: 6.702s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.937 [0.000, 5.000], mean observation: 105.600 [87.000, 236.000], loss: 0.011251, mean_absolute_error: 0.381718, mean_q: -0.442897, mean_eps: 0.100000
223478/500000: episode: 266, duration: 6.955s, episode steps: 783, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 4.662 [0.000, 5.000], mean observation: 105.558 [87.000, 236.000], loss: 0.011163, mean_absolute_error: 0.383064, mean_q: -0.443690, mean_eps: 0.100000
224242/500000: episode: 267, duration: 6.759s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.671 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.011654, mean_absolute_error: 0.386084, mean_q: -0.448935, mean_eps: 0.100000
225006/500000: episode: 268, duration: 6.871s, episode steps: 764, steps per second: 111, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.270 [0.000, 5.000], mean observation: 105.564 [87.000, 236.000], loss: 0.012314, mean_absolute_error: 0.387545, mean_q: -0.449507, mean_eps: 0.100000
225890/500000: episode: 269, duration: 7.803s, episode steps: 884, steps per second: 113, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.023 [0.000, 5.000], mean observation: 105.601 [87.000, 236.000], loss: 0.014232, mean_absolute_error: 0.387409, mean_q: -0.448681, mean_eps: 0.100000
226836/500000: episode: 270, duration: 8.363s, episode steps: 946, steps per second: 113, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 1.127 [0.000, 5.000], mean observation: 105.577 [87.000, 236.000], loss: 0.012875, mean_absolute_error: 0.386557, mean_q: -0.447722, mean_eps: 0.100000
227600/500000: episode: 271, duration: 6.651s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.050 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.014905, mean_absolute_error: 0.387627, mean_q: -0.445722, mean_eps: 0.100000
228544/500000: episode: 272, duration: 8.279s, episode steps: 944, steps per second: 114, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 1.533 [0.000, 5.000], mean observation: 105.593 [87.000, 236.000], loss: 0.012943, mean_absolute_error: 0.386561, mean_q: -0.446845, mean_eps: 0.100000
229368/500000: episode: 273, duration: 7.330s, episode steps: 824, steps per second: 112, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.169 [0.000, 5.000], mean observation: 105.566 [87.000, 236.000], loss: 0.010848, mean_absolute_error: 0.381308, mean_q: -0.443045, mean_eps: 0.100000
230234/500000: episode: 274, duration: 7.607s, episode steps: 866, steps per second: 114, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 3.104 [0.000, 5.000], mean observation: 105.407 [87.000, 236.000], loss: 0.010832, mean_absolute_error: 0.380598, mean_q: -0.444722, mean_eps: 0.100000
231017/500000: episode: 275, duration: 6.856s, episode steps: 783, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.083 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.012768, mean_absolute_error: 0.395622, mean_q: -0.461280, mean_eps: 0.100000
231998/500000: episode: 276, duration: 8.544s, episode steps: 981, steps per second: 115, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action: 1.333 [0.000, 5.000], mean observation: 105.483 [87.000, 236.000], loss: 0.014429, mean_absolute_error: 0.402995, mean_q: -0.465169, mean_eps: 0.100000
232762/500000: episode: 277, duration: 6.786s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.391 [0.000, 5.000], mean observation: 105.636 [87.000, 236.000], loss: 0.011642, mean_absolute_error: 0.402210, mean_q: -0.467039, mean_eps: 0.100000
233605/500000: episode: 278, duration: 7.793s, episode steps: 843, steps per second: 108, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 1.439 [0.000, 5.000], mean observation: 105.559 [87.000, 236.000], loss: 0.013344, mean_absolute_error: 0.402078, mean_q: -0.466879, mean_eps: 0.100000
234397/500000: episode: 279, duration: 6.944s, episode steps: 792, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.064 [0.000, 5.000], mean observation: 105.564 [87.000, 236.000], loss: 0.010750, mean_absolute_error: 0.400415, mean_q: -0.466692, mean_eps: 0.100000
235161/500000: episode: 280, duration: 6.747s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.894 [0.000, 5.000], mean observation: 105.570 [87.000, 236.000], loss: 0.010049, mean_absolute_error: 0.399548, mean_q: -0.465511, mean_eps: 0.100000
235987/500000: episode: 281, duration: 7.215s, episode steps: 826, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.924 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.010935, mean_absolute_error: 0.397351, mean_q: -0.464309, mean_eps: 0.100000
236813/500000: episode: 282, duration: 7.197s, episode steps: 826, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.121 [0.000, 5.000], mean observation: 105.579 [87.000, 236.000], loss: 0.012099, mean_absolute_error: 0.403702, mean_q: -0.470033, mean_eps: 0.100000
237637/500000: episode: 283, duration: 7.186s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.868 [0.000, 5.000], mean observation: 105.586 [87.000, 236.000], loss: 0.013420, mean_absolute_error: 0.401514, mean_q: -0.467081, mean_eps: 0.100000
238584/500000: episode: 284, duration: 8.246s, episode steps: 947, steps per second: 115, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 2.239 [0.000, 5.000], mean observation: 105.452 [87.000, 236.000], loss: 0.013843, mean_absolute_error: 0.407952, mean_q: -0.472459, mean_eps: 0.100000
239348/500000: episode: 285, duration: 6.686s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.906 [0.000, 5.000], mean observation: 105.619 [87.000, 236.000], loss: 0.010830, mean_absolute_error: 0.401904, mean_q: -0.465921, mean_eps: 0.100000
240112/500000: episode: 286, duration: 6.734s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.881 [0.000, 5.000], mean observation: 105.629 [87.000, 236.000], loss: 0.011399, mean_absolute_error: 0.399278, mean_q: -0.464786, mean_eps: 0.100000
240938/500000: episode: 287, duration: 7.191s, episode steps: 826, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.667 [0.000, 5.000], mean observation: 105.628 [87.000, 236.000], loss: 0.012634, mean_absolute_error: 0.410736, mean_q: -0.479481, mean_eps: 0.100000
241702/500000: episode: 288, duration: 6.571s, episode steps: 764, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.204 [0.000, 5.000], mean observation: 105.594 [87.000, 236.000], loss: 0.011903, mean_absolute_error: 0.415893, mean_q: -0.483165, mean_eps: 0.100000
242586/500000: episode: 289, duration: 7.754s, episode steps: 884, steps per second: 114, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.043 [0.000, 5.000], mean observation: 105.581 [87.000, 236.000], loss: 0.012282, mean_absolute_error: 0.418440, mean_q: -0.486662, mean_eps: 0.100000
243567/500000: episode: 290, duration: 8.485s, episode steps: 981, steps per second: 116, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action: 0.234 [0.000, 5.000], mean observation: 105.541 [87.000, 236.000], loss: 0.012048, mean_absolute_error: 0.415758, mean_q: -0.479894, mean_eps: 0.100000
244419/500000: episode: 291, duration: 7.379s, episode steps: 852, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.425 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.012958, mean_absolute_error: 0.416932, mean_q: -0.483611, mean_eps: 0.100000
245183/500000: episode: 292, duration: 6.694s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.292 [0.000, 5.000], mean observation: 105.617 [87.000, 236.000], loss: 0.010332, mean_absolute_error: 0.413458, mean_q: -0.479752, mean_eps: 0.100000
245975/500000: episode: 293, duration: 6.824s, episode steps: 792, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.881 [0.000, 5.000], mean observation: 105.554 [87.000, 236.000], loss: 0.012392, mean_absolute_error: 0.415305, mean_q: -0.482083, mean_eps: 0.100000
246813/500000: episode: 294, duration: 7.208s, episode steps: 838, steps per second: 116, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 4.353 [0.000, 5.000], mean observation: 105.424 [87.000, 236.000], loss: 0.011972, mean_absolute_error: 0.417577, mean_q: -0.484347, mean_eps: 0.100000
247880/500000: episode: 295, duration: 9.277s, episode steps: 1067, steps per second: 115, episode reward: -21.000, mean reward: -0.020 [-1.000, 0.000], mean action: 1.174 [0.000, 5.000], mean observation: 105.605 [87.000, 236.000], loss: 0.012247, mean_absolute_error: 0.417073, mean_q: -0.487344, mean_eps: 0.100000
248760/500000: episode: 296, duration: 7.654s, episode steps: 880, steps per second: 115, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.090 [0.000, 5.000], mean observation: 105.615 [87.000, 236.000], loss: 0.012145, mean_absolute_error: 0.414563, mean_q: -0.484555, mean_eps: 0.100000
249524/500000: episode: 297, duration: 6.619s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.764 [0.000, 5.000], mean observation: 105.566 [87.000, 236.000], loss: 0.012261, mean_absolute_error: 0.416025, mean_q: -0.483235, mean_eps: 0.100000
250307/500000: episode: 298, duration: 6.887s, episode steps: 783, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.484 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.013429, mean_absolute_error: 0.420049, mean_q: -0.488567, mean_eps: 0.100000
251071/500000: episode: 299, duration: 6.719s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.565 [0.000, 5.000], mean observation: 105.584 [87.000, 236.000], loss: 0.013733, mean_absolute_error: 0.435164, mean_q: -0.507356, mean_eps: 0.100000
251909/500000: episode: 300, duration: 7.213s, episode steps: 838, steps per second: 116, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.807 [0.000, 5.000], mean observation: 105.532 [87.000, 236.000], loss: 0.013129, mean_absolute_error: 0.437387, mean_q: -0.510560, mean_eps: 0.100000
252793/500000: episode: 301, duration: 7.622s, episode steps: 884, steps per second: 116, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.256 [0.000, 5.000], mean observation: 105.602 [87.000, 236.000], loss: 0.011791, mean_absolute_error: 0.436766, mean_q: -0.507812, mean_eps: 0.100000
253712/500000: episode: 302, duration: 7.939s, episode steps: 919, steps per second: 116, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 1.395 [0.000, 5.000], mean observation: 105.554 [87.000, 236.000], loss: 0.012174, mean_absolute_error: 0.438089, mean_q: -0.507918, mean_eps: 0.100000
254554/500000: episode: 303, duration: 7.295s, episode steps: 842, steps per second: 115, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.641 [0.000, 5.000], mean observation: 105.535 [87.000, 236.000], loss: 0.012392, mean_absolute_error: 0.435877, mean_q: -0.506061, mean_eps: 0.100000
255378/500000: episode: 304, duration: 7.074s, episode steps: 824, steps per second: 116, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.462 [0.000, 5.000], mean observation: 105.572 [87.000, 236.000], loss: 0.011065, mean_absolute_error: 0.434860, mean_q: -0.505590, mean_eps: 0.100000
256216/500000: episode: 305, duration: 7.302s, episode steps: 838, steps per second: 115, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 3.760 [0.000, 5.000], mean observation: 105.399 [87.000, 236.000], loss: 0.010597, mean_absolute_error: 0.431414, mean_q: -0.505428, mean_eps: 0.100000
257068/500000: episode: 306, duration: 7.417s, episode steps: 852, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.469 [0.000, 5.000], mean observation: 105.592 [87.000, 236.000], loss: 0.010694, mean_absolute_error: 0.432982, mean_q: -0.505731, mean_eps: 0.100000
257832/500000: episode: 307, duration: 6.626s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.496 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.013492, mean_absolute_error: 0.437476, mean_q: -0.510362, mean_eps: 0.100000
258596/500000: episode: 308, duration: 6.652s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.126 [0.000, 5.000], mean observation: 105.581 [87.000, 236.000], loss: 0.013629, mean_absolute_error: 0.437325, mean_q: -0.508306, mean_eps: 0.100000
259420/500000: episode: 309, duration: 7.164s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 4.029 [0.000, 5.000], mean observation: 105.542 [87.000, 236.000], loss: 0.011115, mean_absolute_error: 0.435029, mean_q: -0.506741, mean_eps: 0.100000
260184/500000: episode: 310, duration: 6.947s, episode steps: 764, steps per second: 110, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 4.077 [0.000, 5.000], mean observation: 105.564 [87.000, 236.000], loss: 0.012105, mean_absolute_error: 0.433647, mean_q: -0.506425, mean_eps: 0.100000
260948/500000: episode: 311, duration: 6.554s, episode steps: 764, steps per second: 117, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.327 [0.000, 5.000], mean observation: 105.567 [87.000, 236.000], loss: 0.010850, mean_absolute_error: 0.442375, mean_q: -0.518218, mean_eps: 0.100000
261740/500000: episode: 312, duration: 6.973s, episode steps: 792, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.134 [0.000, 5.000], mean observation: 105.599 [87.000, 236.000], loss: 0.011113, mean_absolute_error: 0.449522, mean_q: -0.525456, mean_eps: 0.100000
262612/500000: episode: 313, duration: 7.709s, episode steps: 872, steps per second: 113, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 1.705 [0.000, 5.000], mean observation: 105.437 [87.000, 236.000], loss: 0.014133, mean_absolute_error: 0.456970, mean_q: -0.530708, mean_eps: 0.100000
263436/500000: episode: 314, duration: 7.241s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 4.177 [0.000, 5.000], mean observation: 105.576 [87.000, 236.000], loss: 0.011211, mean_absolute_error: 0.449946, mean_q: -0.521720, mean_eps: 0.100000
264200/500000: episode: 315, duration: 6.737s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.260 [0.000, 5.000], mean observation: 105.599 [87.000, 236.000], loss: 0.010657, mean_absolute_error: 0.449336, mean_q: -0.522664, mean_eps: 0.100000
264964/500000: episode: 316, duration: 6.651s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.787 [0.000, 5.000], mean observation: 105.589 [87.000, 236.000], loss: 0.012364, mean_absolute_error: 0.450092, mean_q: -0.526750, mean_eps: 0.100000
265848/500000: episode: 317, duration: 7.731s, episode steps: 884, steps per second: 114, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 0.294 [0.000, 5.000], mean observation: 105.616 [87.000, 236.000], loss: 0.010764, mean_absolute_error: 0.451202, mean_q: -0.524728, mean_eps: 0.100000
266788/500000: episode: 318, duration: 8.187s, episode steps: 940, steps per second: 115, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 0.652 [0.000, 5.000], mean observation: 105.644 [87.000, 236.000], loss: 0.011942, mean_absolute_error: 0.452005, mean_q: -0.525249, mean_eps: 0.100000
267552/500000: episode: 319, duration: 6.686s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.170 [0.000, 5.000], mean observation: 105.623 [87.000, 236.000], loss: 0.012584, mean_absolute_error: 0.456847, mean_q: -0.525270, mean_eps: 0.100000
268400/500000: episode: 320, duration: 7.550s, episode steps: 848, steps per second: 112, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.336 [0.000, 5.000], mean observation: 105.625 [87.000, 236.000], loss: 0.013137, mean_absolute_error: 0.455831, mean_q: -0.529460, mean_eps: 0.100000
269284/500000: episode: 321, duration: 7.947s, episode steps: 884, steps per second: 111, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 0.270 [0.000, 5.000], mean observation: 105.605 [87.000, 236.000], loss: 0.011549, mean_absolute_error: 0.454726, mean_q: -0.527027, mean_eps: 0.100000
270108/500000: episode: 322, duration: 7.303s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.606 [0.000, 5.000], mean observation: 105.573 [87.000, 236.000], loss: 0.011550, mean_absolute_error: 0.451277, mean_q: -0.526603, mean_eps: 0.100000
270950/500000: episode: 323, duration: 7.421s, episode steps: 842, steps per second: 113, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 3.081 [0.000, 5.000], mean observation: 105.515 [87.000, 236.000], loss: 0.012190, mean_absolute_error: 0.463207, mean_q: -0.543443, mean_eps: 0.100000
271802/500000: episode: 324, duration: 7.465s, episode steps: 852, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.215 [0.000, 5.000], mean observation: 105.634 [87.000, 236.000], loss: 0.013498, mean_absolute_error: 0.469994, mean_q: -0.549042, mean_eps: 0.100000
272566/500000: episode: 325, duration: 6.617s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.158 [0.000, 5.000], mean observation: 105.600 [87.000, 236.000], loss: 0.010402, mean_absolute_error: 0.466645, mean_q: -0.545523, mean_eps: 0.100000
273330/500000: episode: 326, duration: 6.553s, episode steps: 764, steps per second: 117, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.777 [0.000, 5.000], mean observation: 105.603 [87.000, 236.000], loss: 0.012339, mean_absolute_error: 0.466175, mean_q: -0.546069, mean_eps: 0.100000
274094/500000: episode: 327, duration: 6.605s, episode steps: 764, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.524 [0.000, 5.000], mean observation: 105.616 [87.000, 236.000], loss: 0.011612, mean_absolute_error: 0.467212, mean_q: -0.547765, mean_eps: 0.100000
274999/500000: episode: 328, duration: 7.832s, episode steps: 905, steps per second: 116, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 2.293 [0.000, 5.000], mean observation: 105.587 [87.000, 236.000], loss: 0.011776, mean_absolute_error: 0.468574, mean_q: -0.548284, mean_eps: 0.100000
275823/500000: episode: 329, duration: 7.171s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.342 [0.000, 5.000], mean observation: 105.555 [87.000, 236.000], loss: 0.012328, mean_absolute_error: 0.468637, mean_q: -0.549122, mean_eps: 0.100000
276707/500000: episode: 330, duration: 7.628s, episode steps: 884, steps per second: 116, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 0.270 [0.000, 5.000], mean observation: 105.628 [87.000, 236.000], loss: 0.011359, mean_absolute_error: 0.464314, mean_q: -0.542404, mean_eps: 0.100000
277531/500000: episode: 331, duration: 7.185s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.860 [0.000, 5.000], mean observation: 105.611 [87.000, 236.000], loss: 0.010816, mean_absolute_error: 0.464786, mean_q: -0.543151, mean_eps: 0.100000
278610/500000: episode: 332, duration: 9.381s, episode steps: 1079, steps per second: 115, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action: 0.711 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.012012, mean_absolute_error: 0.467578, mean_q: -0.545452, mean_eps: 0.100000
279480/500000: episode: 333, duration: 7.559s, episode steps: 870, steps per second: 115, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 2.770 [0.000, 5.000], mean observation: 105.496 [87.000, 236.000], loss: 0.011860, mean_absolute_error: 0.468389, mean_q: -0.547270, mean_eps: 0.100000
280263/500000: episode: 334, duration: 6.773s, episode steps: 783, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 4.667 [0.000, 5.000], mean observation: 105.554 [87.000, 236.000], loss: 0.013065, mean_absolute_error: 0.468374, mean_q: -0.546848, mean_eps: 0.100000
281027/500000: episode: 335, duration: 6.630s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.817 [0.000, 5.000], mean observation: 105.569 [87.000, 236.000], loss: 0.011475, mean_absolute_error: 0.476877, mean_q: -0.559334, mean_eps: 0.100000
281819/500000: episode: 336, duration: 6.931s, episode steps: 792, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.986 [0.000, 5.000], mean observation: 105.572 [87.000, 236.000], loss: 0.011988, mean_absolute_error: 0.482101, mean_q: -0.562974, mean_eps: 0.100000
282613/500000: episode: 337, duration: 6.929s, episode steps: 794, steps per second: 115, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 2.626 [0.000, 5.000], mean observation: 105.546 [87.000, 236.000], loss: 0.011900, mean_absolute_error: 0.483653, mean_q: -0.564747, mean_eps: 0.100000
283457/500000: episode: 338, duration: 7.286s, episode steps: 844, steps per second: 116, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 1.468 [0.000, 5.000], mean observation: 105.437 [87.000, 236.000], loss: 0.012775, mean_absolute_error: 0.486379, mean_q: -0.566934, mean_eps: 0.100000
284378/500000: episode: 339, duration: 8.042s, episode steps: 921, steps per second: 115, episode reward: -19.000, mean reward: -0.021 [-1.000, 1.000], mean action: 2.093 [0.000, 5.000], mean observation: 105.474 [87.000, 236.000], loss: 0.012636, mean_absolute_error: 0.485811, mean_q: -0.563844, mean_eps: 0.100000
285220/500000: episode: 340, duration: 7.375s, episode steps: 842, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.612 [0.000, 5.000], mean observation: 105.602 [87.000, 236.000], loss: 0.013832, mean_absolute_error: 0.488778, mean_q: -0.566392, mean_eps: 0.100000
286040/500000: episode: 341, duration: 7.107s, episode steps: 820, steps per second: 115, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 1.493 [0.000, 5.000], mean observation: 105.577 [87.000, 236.000], loss: 0.011171, mean_absolute_error: 0.484564, mean_q: -0.564260, mean_eps: 0.100000
286832/500000: episode: 342, duration: 6.988s, episode steps: 792, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.189 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.011804, mean_absolute_error: 0.486601, mean_q: -0.566637, mean_eps: 0.100000
287596/500000: episode: 343, duration: 6.625s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.785 [0.000, 5.000], mean observation: 105.588 [87.000, 236.000], loss: 0.011079, mean_absolute_error: 0.484855, mean_q: -0.567440, mean_eps: 0.100000
288360/500000: episode: 344, duration: 6.599s, episode steps: 764, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.382 [0.000, 5.000], mean observation: 105.566 [87.000, 236.000], loss: 0.013461, mean_absolute_error: 0.487397, mean_q: -0.569655, mean_eps: 0.100000
289202/500000: episode: 345, duration: 7.396s, episode steps: 842, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 1.709 [0.000, 5.000], mean observation: 105.451 [87.000, 236.000], loss: 0.012065, mean_absolute_error: 0.485810, mean_q: -0.566270, mean_eps: 0.100000
290044/500000: episode: 346, duration: 7.353s, episode steps: 842, steps per second: 115, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 1.284 [0.000, 5.000], mean observation: 105.526 [87.000, 236.000], loss: 0.010827, mean_absolute_error: 0.483023, mean_q: -0.566313, mean_eps: 0.100000
290870/500000: episode: 347, duration: 7.263s, episode steps: 826, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.851 [0.000, 5.000], mean observation: 105.580 [87.000, 236.000], loss: 0.012622, mean_absolute_error: 0.490153, mean_q: -0.574088, mean_eps: 0.100000
291764/500000: episode: 348, duration: 7.790s, episode steps: 894, steps per second: 115, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 1.138 [0.000, 5.000], mean observation: 105.411 [87.000, 236.000], loss: 0.011546, mean_absolute_error: 0.492420, mean_q: -0.575176, mean_eps: 0.100000
292556/500000: episode: 349, duration: 6.874s, episode steps: 792, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.059 [0.000, 5.000], mean observation: 105.540 [87.000, 236.000], loss: 0.012720, mean_absolute_error: 0.494766, mean_q: -0.579085, mean_eps: 0.100000
293320/500000: episode: 350, duration: 6.705s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.582 [0.000, 5.000], mean observation: 105.578 [87.000, 236.000], loss: 0.011843, mean_absolute_error: 0.493220, mean_q: -0.578204, mean_eps: 0.100000
294144/500000: episode: 351, duration: 7.123s, episode steps: 824, steps per second: 116, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.807 [0.000, 5.000], mean observation: 105.578 [87.000, 236.000], loss: 0.012439, mean_absolute_error: 0.494070, mean_q: -0.578063, mean_eps: 0.100000
294908/500000: episode: 352, duration: 6.984s, episode steps: 764, steps per second: 109, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.499 [0.000, 5.000], mean observation: 105.581 [87.000, 236.000], loss: 0.011808, mean_absolute_error: 0.493894, mean_q: -0.577936, mean_eps: 0.100000
295732/500000: episode: 353, duration: 7.209s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.374 [0.000, 5.000], mean observation: 105.586 [87.000, 236.000], loss: 0.012938, mean_absolute_error: 0.496884, mean_q: -0.578414, mean_eps: 0.100000
296524/500000: episode: 354, duration: 6.810s, episode steps: 792, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.604 [0.000, 5.000], mean observation: 105.569 [87.000, 236.000], loss: 0.011634, mean_absolute_error: 0.494967, mean_q: -0.576962, mean_eps: 0.100000
297288/500000: episode: 355, duration: 6.624s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.310 [0.000, 5.000], mean observation: 105.584 [87.000, 236.000], loss: 0.012649, mean_absolute_error: 0.494079, mean_q: -0.577449, mean_eps: 0.100000
298114/500000: episode: 356, duration: 7.166s, episode steps: 826, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.338 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.012754, mean_absolute_error: 0.495620, mean_q: -0.580713, mean_eps: 0.100000
298956/500000: episode: 357, duration: 7.350s, episode steps: 842, steps per second: 115, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 1.158 [0.000, 5.000], mean observation: 105.497 [87.000, 236.000], loss: 0.012228, mean_absolute_error: 0.495641, mean_q: -0.574264, mean_eps: 0.100000
299720/500000: episode: 358, duration: 6.736s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.819 [0.000, 5.000], mean observation: 105.584 [87.000, 236.000], loss: 0.012964, mean_absolute_error: 0.496095, mean_q: -0.578857, mean_eps: 0.100000
300546/500000: episode: 359, duration: 7.138s, episode steps: 826, steps per second: 116, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.923 [0.000, 5.000], mean observation: 105.541 [87.000, 236.000], loss: 0.012515, mean_absolute_error: 0.498651, mean_q: -0.583811, mean_eps: 0.100000
301460/500000: episode: 360, duration: 7.981s, episode steps: 914, steps per second: 115, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 1.460 [0.000, 5.000], mean observation: 105.602 [87.000, 236.000], loss: 0.012562, mean_absolute_error: 0.511472, mean_q: -0.597671, mean_eps: 0.100000
302397/500000: episode: 361, duration: 8.257s, episode steps: 937, steps per second: 113, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action: 0.896 [0.000, 5.000], mean observation: 105.501 [87.000, 236.000], loss: 0.013885, mean_absolute_error: 0.516076, mean_q: -0.600613, mean_eps: 0.100000
303221/500000: episode: 362, duration: 7.164s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.886 [0.000, 5.000], mean observation: 105.645 [87.000, 236.000], loss: 0.011941, mean_absolute_error: 0.514203, mean_q: -0.599982, mean_eps: 0.100000
303985/500000: episode: 363, duration: 6.934s, episode steps: 764, steps per second: 110, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.914 [0.000, 5.000], mean observation: 105.600 [87.000, 236.000], loss: 0.011561, mean_absolute_error: 0.507930, mean_q: -0.595007, mean_eps: 0.100000
304777/500000: episode: 364, duration: 7.105s, episode steps: 792, steps per second: 111, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.375 [0.000, 5.000], mean observation: 105.583 [87.000, 236.000], loss: 0.011016, mean_absolute_error: 0.509600, mean_q: -0.598537, mean_eps: 0.100000
305541/500000: episode: 365, duration: 6.662s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.124 [0.000, 5.000], mean observation: 105.569 [87.000, 236.000], loss: 0.012192, mean_absolute_error: 0.509522, mean_q: -0.595621, mean_eps: 0.100000
306367/500000: episode: 366, duration: 7.150s, episode steps: 826, steps per second: 116, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.960 [0.000, 5.000], mean observation: 105.541 [87.000, 236.000], loss: 0.012444, mean_absolute_error: 0.509949, mean_q: -0.593987, mean_eps: 0.100000
307131/500000: episode: 367, duration: 6.630s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.580 [0.000, 5.000], mean observation: 105.588 [87.000, 236.000], loss: 0.013306, mean_absolute_error: 0.514000, mean_q: -0.601512, mean_eps: 0.100000
307951/500000: episode: 368, duration: 7.126s, episode steps: 820, steps per second: 115, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 2.393 [0.000, 5.000], mean observation: 105.564 [87.000, 236.000], loss: 0.012054, mean_absolute_error: 0.511546, mean_q: -0.600340, mean_eps: 0.100000
308799/500000: episode: 369, duration: 7.389s, episode steps: 848, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.438 [0.000, 5.000], mean observation: 105.562 [87.000, 236.000], loss: 0.011910, mean_absolute_error: 0.513059, mean_q: -0.599439, mean_eps: 0.100000
309563/500000: episode: 370, duration: 6.631s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.817 [0.000, 5.000], mean observation: 105.604 [87.000, 236.000], loss: 0.013258, mean_absolute_error: 0.518032, mean_q: -0.602629, mean_eps: 0.100000
310447/500000: episode: 371, duration: 7.716s, episode steps: 884, steps per second: 115, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 1.751 [0.000, 5.000], mean observation: 105.561 [87.000, 236.000], loss: 0.012178, mean_absolute_error: 0.518167, mean_q: -0.606722, mean_eps: 0.100000
311211/500000: episode: 372, duration: 6.677s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.749 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.011810, mean_absolute_error: 0.529420, mean_q: -0.617235, mean_eps: 0.100000
311975/500000: episode: 373, duration: 6.693s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.877 [0.000, 5.000], mean observation: 105.571 [87.000, 236.000], loss: 0.010810, mean_absolute_error: 0.527537, mean_q: -0.616040, mean_eps: 0.100000
312859/500000: episode: 374, duration: 7.657s, episode steps: 884, steps per second: 115, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 1.045 [0.000, 5.000], mean observation: 105.591 [87.000, 236.000], loss: 0.012661, mean_absolute_error: 0.528355, mean_q: -0.617041, mean_eps: 0.100000
313701/500000: episode: 375, duration: 7.347s, episode steps: 842, steps per second: 115, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.596 [0.000, 5.000], mean observation: 105.382 [87.000, 236.000], loss: 0.012835, mean_absolute_error: 0.531445, mean_q: -0.619147, mean_eps: 0.100000
314539/500000: episode: 376, duration: 7.311s, episode steps: 838, steps per second: 115, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 4.421 [0.000, 5.000], mean observation: 105.397 [87.000, 236.000], loss: 0.014203, mean_absolute_error: 0.534328, mean_q: -0.620138, mean_eps: 0.100000
315377/500000: episode: 377, duration: 7.362s, episode steps: 838, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 4.732 [0.000, 5.000], mean observation: 105.420 [87.000, 236.000], loss: 0.011772, mean_absolute_error: 0.534095, mean_q: -0.621231, mean_eps: 0.100000
316201/500000: episode: 378, duration: 7.222s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.331 [0.000, 5.000], mean observation: 105.578 [87.000, 236.000], loss: 0.013119, mean_absolute_error: 0.533871, mean_q: -0.619568, mean_eps: 0.100000
317025/500000: episode: 379, duration: 7.220s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.598 [0.000, 5.000], mean observation: 105.591 [87.000, 236.000], loss: 0.013007, mean_absolute_error: 0.530198, mean_q: -0.614415, mean_eps: 0.100000
317877/500000: episode: 380, duration: 7.449s, episode steps: 852, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.846 [0.000, 5.000], mean observation: 105.632 [87.000, 236.000], loss: 0.011887, mean_absolute_error: 0.530428, mean_q: -0.614399, mean_eps: 0.100000
318641/500000: episode: 381, duration: 6.669s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.262 [0.000, 5.000], mean observation: 105.587 [87.000, 236.000], loss: 0.011867, mean_absolute_error: 0.532404, mean_q: -0.619697, mean_eps: 0.100000
319479/500000: episode: 382, duration: 7.252s, episode steps: 838, steps per second: 116, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 4.166 [0.000, 5.000], mean observation: 105.515 [87.000, 236.000], loss: 0.012794, mean_absolute_error: 0.532994, mean_q: -0.623046, mean_eps: 0.100000
320243/500000: episode: 383, duration: 6.654s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.128 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.010828, mean_absolute_error: 0.530235, mean_q: -0.622389, mean_eps: 0.100000
321007/500000: episode: 384, duration: 6.585s, episode steps: 764, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.937 [0.000, 5.000], mean observation: 105.567 [87.000, 236.000], loss: 0.014019, mean_absolute_error: 0.540115, mean_q: -0.630754, mean_eps: 0.100000
321771/500000: episode: 385, duration: 6.639s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.474 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.012619, mean_absolute_error: 0.542896, mean_q: -0.634487, mean_eps: 0.100000
322535/500000: episode: 386, duration: 6.608s, episode steps: 764, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.436 [0.000, 5.000], mean observation: 105.583 [87.000, 236.000], loss: 0.014707, mean_absolute_error: 0.546612, mean_q: -0.639236, mean_eps: 0.100000
323299/500000: episode: 387, duration: 6.587s, episode steps: 764, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.518 [0.000, 5.000], mean observation: 105.613 [87.000, 236.000], loss: 0.011621, mean_absolute_error: 0.544099, mean_q: -0.635188, mean_eps: 0.100000
324123/500000: episode: 388, duration: 7.238s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.848 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.013102, mean_absolute_error: 0.544709, mean_q: -0.634807, mean_eps: 0.100000
324965/500000: episode: 389, duration: 7.319s, episode steps: 842, steps per second: 115, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 1.588 [0.000, 5.000], mean observation: 105.445 [87.000, 236.000], loss: 0.012248, mean_absolute_error: 0.544949, mean_q: -0.635222, mean_eps: 0.100000
325803/500000: episode: 390, duration: 7.295s, episode steps: 838, steps per second: 115, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 3.977 [0.000, 5.000], mean observation: 105.531 [87.000, 236.000], loss: 0.011113, mean_absolute_error: 0.539816, mean_q: -0.633119, mean_eps: 0.100000
326681/500000: episode: 391, duration: 7.660s, episode steps: 878, steps per second: 115, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 1.411 [0.000, 5.000], mean observation: 105.439 [87.000, 236.000], loss: 0.013013, mean_absolute_error: 0.546939, mean_q: -0.640082, mean_eps: 0.100000
327445/500000: episode: 392, duration: 6.622s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.950 [0.000, 5.000], mean observation: 105.604 [87.000, 236.000], loss: 0.013642, mean_absolute_error: 0.545325, mean_q: -0.635578, mean_eps: 0.100000
328209/500000: episode: 393, duration: 6.655s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.376 [0.000, 5.000], mean observation: 105.619 [87.000, 236.000], loss: 0.011597, mean_absolute_error: 0.544256, mean_q: -0.637632, mean_eps: 0.100000
328973/500000: episode: 394, duration: 6.646s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.753 [0.000, 5.000], mean observation: 105.586 [87.000, 236.000], loss: 0.011752, mean_absolute_error: 0.541464, mean_q: -0.632949, mean_eps: 0.100000
329737/500000: episode: 395, duration: 6.772s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.866 [0.000, 5.000], mean observation: 105.579 [87.000, 236.000], loss: 0.013720, mean_absolute_error: 0.543640, mean_q: -0.636478, mean_eps: 0.100000
330501/500000: episode: 396, duration: 6.705s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.741 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.011846, mean_absolute_error: 0.546088, mean_q: -0.640388, mean_eps: 0.100000
331265/500000: episode: 397, duration: 6.653s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.486 [0.000, 5.000], mean observation: 105.571 [87.000, 236.000], loss: 0.012120, mean_absolute_error: 0.555671, mean_q: -0.650403, mean_eps: 0.100000
332029/500000: episode: 398, duration: 6.679s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.872 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.013151, mean_absolute_error: 0.566645, mean_q: -0.664026, mean_eps: 0.100000
332812/500000: episode: 399, duration: 6.866s, episode steps: 783, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.978 [0.000, 5.000], mean observation: 105.571 [87.000, 236.000], loss: 0.012407, mean_absolute_error: 0.564261, mean_q: -0.659233, mean_eps: 0.100000
333576/500000: episode: 400, duration: 6.672s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.551 [0.000, 5.000], mean observation: 105.567 [87.000, 236.000], loss: 0.011837, mean_absolute_error: 0.562773, mean_q: -0.657125, mean_eps: 0.100000
334368/500000: episode: 401, duration: 7.040s, episode steps: 792, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 4.715 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.015099, mean_absolute_error: 0.567759, mean_q: -0.662903, mean_eps: 0.100000
335132/500000: episode: 402, duration: 6.738s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.853 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.013877, mean_absolute_error: 0.565570, mean_q: -0.661240, mean_eps: 0.100000
335896/500000: episode: 403, duration: 6.700s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.531 [0.000, 5.000], mean observation: 105.573 [87.000, 236.000], loss: 0.012680, mean_absolute_error: 0.560494, mean_q: -0.656505, mean_eps: 0.100000
336679/500000: episode: 404, duration: 6.919s, episode steps: 783, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.556 [0.000, 5.000], mean observation: 105.544 [87.000, 236.000], loss: 0.012523, mean_absolute_error: 0.560128, mean_q: -0.656338, mean_eps: 0.100000
337503/500000: episode: 405, duration: 7.214s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.598 [0.000, 5.000], mean observation: 105.569 [87.000, 236.000], loss: 0.011585, mean_absolute_error: 0.557135, mean_q: -0.653058, mean_eps: 0.100000
338405/500000: episode: 406, duration: 7.837s, episode steps: 902, steps per second: 115, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 1.623 [0.000, 5.000], mean observation: 105.580 [87.000, 236.000], loss: 0.012481, mean_absolute_error: 0.555532, mean_q: -0.652971, mean_eps: 0.100000
339229/500000: episode: 407, duration: 7.327s, episode steps: 824, steps per second: 112, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.530 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.010569, mean_absolute_error: 0.552827, mean_q: -0.649048, mean_eps: 0.100000
340012/500000: episode: 408, duration: 7.072s, episode steps: 783, steps per second: 111, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 4.215 [0.000, 5.000], mean observation: 105.570 [87.000, 236.000], loss: 0.011075, mean_absolute_error: 0.557623, mean_q: -0.654789, mean_eps: 0.100000
340804/500000: episode: 409, duration: 6.947s, episode steps: 792, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.309 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.012220, mean_absolute_error: 0.571389, mean_q: -0.671823, mean_eps: 0.100000
341642/500000: episode: 410, duration: 7.360s, episode steps: 838, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.760 [0.000, 5.000], mean observation: 105.494 [87.000, 236.000], loss: 0.011385, mean_absolute_error: 0.573201, mean_q: -0.674099, mean_eps: 0.100000
342406/500000: episode: 411, duration: 6.721s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.010 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.011933, mean_absolute_error: 0.574903, mean_q: -0.676941, mean_eps: 0.100000
343170/500000: episode: 412, duration: 6.749s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.102 [0.000, 5.000], mean observation: 105.569 [87.000, 236.000], loss: 0.013050, mean_absolute_error: 0.578335, mean_q: -0.675316, mean_eps: 0.100000
343934/500000: episode: 413, duration: 6.666s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.438 [0.000, 5.000], mean observation: 105.575 [87.000, 236.000], loss: 0.014039, mean_absolute_error: 0.584132, mean_q: -0.680588, mean_eps: 0.100000
344772/500000: episode: 414, duration: 7.361s, episode steps: 838, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 4.246 [0.000, 5.000], mean observation: 105.412 [87.000, 236.000], loss: 0.011681, mean_absolute_error: 0.576254, mean_q: -0.675212, mean_eps: 0.100000
345596/500000: episode: 415, duration: 7.219s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.819 [0.000, 5.000], mean observation: 105.611 [87.000, 236.000], loss: 0.014036, mean_absolute_error: 0.580184, mean_q: -0.679378, mean_eps: 0.100000
346360/500000: episode: 416, duration: 6.664s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.649 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.011472, mean_absolute_error: 0.579061, mean_q: -0.679427, mean_eps: 0.100000
347212/500000: episode: 417, duration: 7.441s, episode steps: 852, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.516 [0.000, 5.000], mean observation: 105.581 [87.000, 236.000], loss: 0.011808, mean_absolute_error: 0.577525, mean_q: -0.677084, mean_eps: 0.100000
348037/500000: episode: 418, duration: 7.276s, episode steps: 825, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.931 [0.000, 5.000], mean observation: 105.542 [87.000, 236.000], loss: 0.014161, mean_absolute_error: 0.582295, mean_q: -0.681229, mean_eps: 0.100000
348801/500000: episode: 419, duration: 6.697s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.658 [0.000, 5.000], mean observation: 105.567 [87.000, 236.000], loss: 0.011356, mean_absolute_error: 0.576712, mean_q: -0.677367, mean_eps: 0.100000
349720/500000: episode: 420, duration: 8.050s, episode steps: 919, steps per second: 114, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 3.177 [0.000, 5.000], mean observation: 105.498 [87.000, 236.000], loss: 0.012159, mean_absolute_error: 0.574766, mean_q: -0.673969, mean_eps: 0.100000
350562/500000: episode: 421, duration: 7.391s, episode steps: 842, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.738 [0.000, 5.000], mean observation: 105.490 [87.000, 236.000], loss: 0.011128, mean_absolute_error: 0.579124, mean_q: -0.680857, mean_eps: 0.100000
351354/500000: episode: 422, duration: 6.929s, episode steps: 792, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.230 [0.000, 5.000], mean observation: 105.559 [87.000, 236.000], loss: 0.012538, mean_absolute_error: 0.585912, mean_q: -0.690183, mean_eps: 0.100000
352118/500000: episode: 423, duration: 6.685s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 4.480 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.009719, mean_absolute_error: 0.590252, mean_q: -0.695538, mean_eps: 0.100000
352882/500000: episode: 424, duration: 6.726s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.649 [0.000, 5.000], mean observation: 105.566 [87.000, 236.000], loss: 0.013100, mean_absolute_error: 0.590278, mean_q: -0.691394, mean_eps: 0.100000
353706/500000: episode: 425, duration: 7.140s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.403 [0.000, 5.000], mean observation: 105.594 [87.000, 236.000], loss: 0.012058, mean_absolute_error: 0.591947, mean_q: -0.692448, mean_eps: 0.100000
354470/500000: episode: 426, duration: 6.692s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.170 [0.000, 5.000], mean observation: 105.583 [87.000, 236.000], loss: 0.011167, mean_absolute_error: 0.588189, mean_q: -0.691306, mean_eps: 0.100000
355314/500000: episode: 427, duration: 7.348s, episode steps: 844, steps per second: 115, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 0.930 [0.000, 5.000], mean observation: 105.445 [87.000, 236.000], loss: 0.011937, mean_absolute_error: 0.587815, mean_q: -0.690519, mean_eps: 0.100000
356078/500000: episode: 428, duration: 6.745s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.818 [0.000, 5.000], mean observation: 105.594 [87.000, 236.000], loss: 0.012571, mean_absolute_error: 0.590855, mean_q: -0.692097, mean_eps: 0.100000
```

```
356999/500000: episode: 429, duration: 7.986s, episode steps: 921, steps per second: 115, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 0.712 [0.000, 5.000], mean observation: 105.404 [87.000, 236.000], loss: 0.013786, mean_absolute_error: 0.596760, mean_q: -0.698799, mean_eps: 0.100000
357823/500000: episode: 430, duration: 7.187s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.681 [0.000, 5.000], mean observation: 105.600 [87.000, 236.000], loss: 0.013364, mean_absolute_error: 0.593708, mean_q: -0.699183, mean_eps: 0.100000
358587/500000: episode: 431, duration: 6.702s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.315 [0.000, 5.000], mean observation: 105.580 [87.000, 236.000], loss: 0.013897, mean_absolute_error: 0.598067, mean_q: -0.701501, mean_eps: 0.100000
359351/500000: episode: 432, duration: 6.646s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.673 [0.000, 5.000], mean observation: 105.619 [87.000, 236.000], loss: 0.013278, mean_absolute_error: 0.596466, mean_q: -0.697393, mean_eps: 0.100000
360115/500000: episode: 433, duration: 6.664s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.064 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.012673, mean_absolute_error: 0.594377, mean_q: -0.694229, mean_eps: 0.100000
360879/500000: episode: 434, duration: 6.641s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.683 [0.000, 5.000], mean observation: 105.572 [87.000, 236.000], loss: 0.012273, mean_absolute_error: 0.601453, mean_q: -0.705229, mean_eps: 0.100000
361643/500000: episode: 435, duration: 6.715s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.359 [0.000, 5.000], mean observation: 105.566 [87.000, 236.000], loss: 0.010878, mean_absolute_error: 0.603883, mean_q: -0.709706, mean_eps: 0.100000
362435/500000: episode: 436, duration: 6.858s, episode steps: 792, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.842 [0.000, 5.000], mean observation: 105.564 [87.000, 236.000], loss: 0.011920, mean_absolute_error: 0.604411, mean_q: -0.712417, mean_eps: 0.100000
363356/500000: episode: 437, duration: 8.065s, episode steps: 921, steps per second: 114, episode reward: -19.000, mean reward: -0.021 [-1.000, 1.000], mean action: 1.598 [0.000, 5.000], mean observation: 105.505 [87.000, 236.000], loss: 0.014529, mean_absolute_error: 0.613857, mean_q: -0.719178, mean_eps: 0.100000
364180/500000: episode: 438, duration: 7.338s, episode steps: 824, steps per second: 112, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.657 [0.000, 5.000], mean observation: 105.602 [87.000, 236.000], loss: 0.012355, mean_absolute_error: 0.614925, mean_q: -0.723080, mean_eps: 0.100000
364972/500000: episode: 439, duration: 7.124s, episode steps: 792, steps per second: 111, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.898 [0.000, 5.000], mean observation: 105.602 [87.000, 236.000], loss: 0.013040, mean_absolute_error: 0.615960, mean_q: -0.722579, mean_eps: 0.100000
365798/500000: episode: 440, duration: 7.182s, episode steps: 826, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.656 [0.000, 5.000], mean observation: 105.598 [87.000, 236.000], loss: 0.011400, mean_absolute_error: 0.610064, mean_q: -0.716793, mean_eps: 0.100000
366562/500000: episode: 441, duration: 6.663s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.119 [0.000, 5.000], mean observation: 105.597 [87.000, 236.000], loss: 0.011291, mean_absolute_error: 0.607879, mean_q: -0.715524, mean_eps: 0.100000
367326/500000: episode: 442, duration: 6.779s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.975 [0.000, 5.000], mean observation: 105.578 [87.000, 236.000], loss: 0.012695, mean_absolute_error: 0.613396, mean_q: -0.719415, mean_eps: 0.100000
368090/500000: episode: 443, duration: 6.617s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.838 [0.000, 5.000], mean observation: 105.564 [87.000, 236.000], loss: 0.013190, mean_absolute_error: 0.618405, mean_q: -0.724708, mean_eps: 0.100000
368902/500000: episode: 444, duration: 7.074s, episode steps: 812, steps per second: 115, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 2.618 [0.000, 5.000], mean observation: 105.596 [87.000, 236.000], loss: 0.012558, mean_absolute_error: 0.612181, mean_q: -0.721094, mean_eps: 0.100000
369666/500000: episode: 445, duration: 6.578s, episode steps: 764, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.110 [0.000, 5.000], mean observation: 105.572 [87.000, 236.000], loss: 0.011987, mean_absolute_error: 0.611546, mean_q: -0.719241, mean_eps: 0.100000
370492/500000: episode: 446, duration: 7.233s, episode steps: 826, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.875 [0.000, 5.000], mean observation: 105.577 [87.000, 236.000], loss: 0.013742, mean_absolute_error: 0.613209, mean_q: -0.720785, mean_eps: 0.100000
371400/500000: episode: 447, duration: 7.964s, episode steps: 908, steps per second: 114, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 1.716 [0.000, 5.000], mean observation: 105.580 [87.000, 236.000], loss: 0.012484, mean_absolute_error: 0.625616, mean_q: -0.736203, mean_eps: 0.100000
372193/500000: episode: 448, duration: 6.934s, episode steps: 793, steps per second: 114, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 1.816 [0.000, 5.000], mean observation: 105.576 [87.000, 236.000], loss: 0.012580, mean_absolute_error: 0.630317, mean_q: -0.740146, mean_eps: 0.100000
372957/500000: episode: 449, duration: 6.596s, episode steps: 764, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.683 [0.000, 5.000], mean observation: 105.570 [87.000, 236.000], loss: 0.011757, mean_absolute_error: 0.624632, mean_q: -0.735095, mean_eps: 0.100000
373721/500000: episode: 450, duration: 6.746s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.709 [0.000, 5.000], mean observation: 105.612 [87.000, 236.000], loss: 0.012471, mean_absolute_error: 0.630233, mean_q: -0.740685, mean_eps: 0.100000
374545/500000: episode: 451, duration: 7.278s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.046 [0.000, 5.000], mean observation: 105.538 [87.000, 236.000], loss: 0.012614, mean_absolute_error: 0.626038, mean_q: -0.732940, mean_eps: 0.100000
375369/500000: episode: 452, duration: 7.559s, episode steps: 824, steps per second: 109, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.756 [0.000, 5.000], mean observation: 105.554 [87.000, 236.000], loss: 0.010463, mean_absolute_error: 0.622491, mean_q: -0.731807, mean_eps: 0.100000
376133/500000: episode: 453, duration: 6.789s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.500 [0.000, 5.000], mean observation: 105.604 [87.000, 236.000], loss: 0.013268, mean_absolute_error: 0.621910, mean_q: -0.728576, mean_eps: 0.100000
377019/500000: episode: 454, duration: 7.900s, episode steps: 886, steps per second: 112, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 1.663 [0.000, 5.000], mean observation: 105.586 [87.000, 236.000], loss: 0.011908, mean_absolute_error: 0.622925, mean_q: -0.729778, mean_eps: 0.100000
377843/500000: episode: 455, duration: 7.321s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.945 [0.000, 5.000], mean observation: 105.593 [87.000, 236.000], loss: 0.012170, mean_absolute_error: 0.622846, mean_q: -0.728789, mean_eps: 0.100000
378607/500000: episode: 456, duration: 6.591s, episode steps: 764, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.271 [0.000, 5.000], mean observation: 105.570 [87.000, 236.000], loss: 0.011485, mean_absolute_error: 0.623289, mean_q: -0.734555, mean_eps: 0.100000
379371/500000: episode: 457, duration: 6.690s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.526 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.012935, mean_absolute_error: 0.627148, mean_q: -0.735713, mean_eps: 0.100000
380135/500000: episode: 458, duration: 6.658s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.309 [0.000, 5.000], mean observation: 105.606 [87.000, 236.000], loss: 0.011248, mean_absolute_error: 0.621315, mean_q: -0.730809, mean_eps: 0.100000
380899/500000: episode: 459, duration: 6.717s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.897 [0.000, 5.000], mean observation: 105.588 [87.000, 236.000], loss: 0.013943, mean_absolute_error: 0.633476, mean_q: -0.743250, mean_eps: 0.100000
381663/500000: episode: 460, duration: 6.608s, episode steps: 764, steps per second: 116, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.842 [0.000, 5.000], mean observation: 105.567 [87.000, 236.000], loss: 0.011222, mean_absolute_error: 0.637340, mean_q: -0.752132, mean_eps: 0.100000
382427/500000: episode: 461, duration: 6.681s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.191 [0.000, 5.000], mean observation: 105.590 [87.000, 236.000], loss: 0.011873, mean_absolute_error: 0.637447, mean_q: -0.747566, mean_eps: 0.100000
383253/500000: episode: 462, duration: 7.330s, episode steps: 826, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.587 [0.000, 5.000], mean observation: 105.573 [87.000, 236.000], loss: 0.012641, mean_absolute_error: 0.638876, mean_q: -0.749656, mean_eps: 0.100000
384017/500000: episode: 463, duration: 6.699s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.305 [0.000, 5.000], mean observation: 105.594 [87.000, 236.000], loss: 0.012331, mean_absolute_error: 0.637657, mean_q: -0.750203, mean_eps: 0.100000
384865/500000: episode: 464, duration: 7.521s, episode steps: 848, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.180 [0.000, 5.000], mean observation: 105.609 [87.000, 236.000], loss: 0.013060, mean_absolute_error: 0.636273, mean_q: -0.746601, mean_eps: 0.100000
385629/500000: episode: 465, duration: 6.733s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.445 [0.000, 5.000], mean observation: 105.582 [87.000, 236.000], loss: 0.013521, mean_absolute_error: 0.642696, mean_q: -0.753275, mean_eps: 0.100000
386471/500000: episode: 466, duration: 7.387s, episode steps: 842, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 1.474 [0.000, 5.000], mean observation: 105.533 [87.000, 236.000], loss: 0.011892, mean_absolute_error: 0.639675, mean_q: -0.750783, mean_eps: 0.100000
387235/500000: episode: 467, duration: 6.824s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.338 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.012081, mean_absolute_error: 0.635091, mean_q: -0.746863, mean_eps: 0.100000
387999/500000: episode: 468, duration: 6.706s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.788 [0.000, 5.000], mean observation: 105.574 [87.000, 236.000], loss: 0.012030, mean_absolute_error: 0.635883, mean_q: -0.747498, mean_eps: 0.100000
388791/500000: episode: 469, duration: 6.977s, episode steps: 792, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.396 [0.000, 5.000], mean observation: 105.579 [87.000, 236.000], loss: 0.013028, mean_absolute_error: 0.637841, mean_q: -0.751629, mean_eps: 0.100000
389584/500000: episode: 470, duration: 6.899s, episode steps: 793, steps per second: 115, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 1.113 [0.000, 5.000], mean observation: 105.627 [87.000, 236.000], loss: 0.011052, mean_absolute_error: 0.635545, mean_q: -0.746161, mean_eps: 0.100000
390465/500000: episode: 471, duration: 7.733s, episode steps: 881, steps per second: 114, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 0.740 [0.000, 5.000], mean observation: 105.619 [87.000, 236.000], loss: 0.012826, mean_absolute_error: 0.638146, mean_q: -0.749950, mean_eps: 0.100000
391257/500000: episode: 472, duration: 6.860s, episode steps: 792, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.058 [0.000, 5.000], mean observation: 105.551 [87.000, 236.000], loss: 0.011155, mean_absolute_error: 0.644936, mean_q: -0.761757, mean_eps: 0.100000
392021/500000: episode: 473, duration: 6.728s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.441 [0.000, 5.000], mean observation: 105.578 [87.000, 236.000], loss: 0.012410, mean_absolute_error: 0.648017, mean_q: -0.763015, mean_eps: 0.100000
392815/500000: episode: 474, duration: 6.960s, episode steps: 794, steps per second: 114, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 1.519 [0.000, 5.000], mean observation: 105.610 [87.000, 236.000], loss: 0.013553, mean_absolute_error: 0.650020, mean_q: -0.764347, mean_eps: 0.100000
393607/500000: episode: 475, duration: 6.914s, episode steps: 792, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.138 [0.000, 5.000], mean observation: 105.566 [87.000, 236.000], loss: 0.012767, mean_absolute_error: 0.649989, mean_q: -0.763653, mean_eps: 0.100000
394371/500000: episode: 476, duration: 6.623s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.686 [0.000, 5.000], mean observation: 105.599 [87.000, 236.000], loss: 0.010482, mean_absolute_error: 0.646447, mean_q: -0.762404, mean_eps: 0.100000
395135/500000: episode: 477, duration: 6.749s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.825 [0.000, 5.000], mean observation: 105.570 [87.000, 236.000], loss: 0.011556, mean_absolute_error: 0.649584, mean_q: -0.763239, mean_eps: 0.100000
395899/500000: episode: 478, duration: 6.653s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.777 [0.000, 5.000], mean observation: 105.572 [87.000, 236.000], loss: 0.012567, mean_absolute_error: 0.654199, mean_q: -0.770225, mean_eps: 0.100000
396691/500000: episode: 479, duration: 6.894s, episode steps: 792, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.234 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.012655, mean_absolute_error: 0.650355, mean_q: -0.764922, mean_eps: 0.100000
397773/500000: episode: 480, duration: 9.461s, episode steps: 1082, steps per second: 114, episode reward: -20.000, mean reward: -0.018 [-1.000, 1.000], mean action: 1.439 [0.000, 5.000], mean observation: 105.472 [87.000, 236.000], loss: 0.012755, mean_absolute_error: 0.651136, mean_q: -0.764426, mean_eps: 0.100000
398597/500000: episode: 481, duration: 7.228s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.234 [0.000, 5.000], mean observation: 105.625 [87.000, 236.000], loss: 0.012674, mean_absolute_error: 0.646191, mean_q: -0.756117, mean_eps: 0.100000
399361/500000: episode: 482, duration: 7.039s, episode steps: 764, steps per second: 109, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.238 [0.000, 5.000], mean observation: 105.576 [87.000, 236.000], loss: 0.012922, mean_absolute_error: 0.647574, mean_q: -0.759988, mean_eps: 0.100000
400125/500000: episode: 483, duration: 6.722s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.103 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.012408, mean_absolute_error: 0.652521, mean_q: -0.767568, mean_eps: 0.100000
400949/500000: episode: 484, duration: 7.233s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.544 [0.000, 5.000], mean observation: 105.576 [87.000, 236.000], loss: 0.012466, mean_absolute_error: 0.665469, mean_q: -0.782666, mean_eps: 0.100000
401773/500000: episode: 485, duration: 7.211s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.522 [0.000, 5.000], mean observation: 105.580 [87.000, 236.000], loss: 0.011725, mean_absolute_error: 0.666173, mean_q: -0.781857, mean_eps: 0.100000
402814/500000: episode: 486, duration: 9.074s, episode steps: 1041, steps per second: 115, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action: 1.601 [0.000, 5.000], mean observation: 105.542 [87.000, 236.000], loss: 0.012528, mean_absolute_error: 0.666909, mean_q: -0.783092, mean_eps: 0.100000
403578/500000: episode: 487, duration: 6.629s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.128 [0.000, 5.000], mean observation: 105.566 [87.000, 236.000], loss: 0.012630, mean_absolute_error: 0.667477, mean_q: -0.785950, mean_eps: 0.100000
404462/500000: episode: 488, duration: 7.735s, episode steps: 884, steps per second: 114, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 0.811 [0.000, 5.000], mean observation: 105.575 [87.000, 236.000], loss: 0.011871, mean_absolute_error: 0.667389, mean_q: -0.784805, mean_eps: 0.100000
405226/500000: episode: 489, duration: 6.701s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.699 [0.000, 5.000], mean observation: 105.585 [87.000, 236.000], loss: 0.010223, mean_absolute_error: 0.665524, mean_q: -0.787331, mean_eps: 0.100000
405990/500000: episode: 490, duration: 6.634s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.551 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.011161, mean_absolute_error: 0.664623, mean_q: -0.784627, mean_eps: 0.100000
406772/500000: episode: 491, duration: 6.789s, episode steps: 782, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.972 [0.000, 5.000], mean observation: 105.556 [87.000, 236.000], loss: 0.011477, mean_absolute_error: 0.667096, mean_q: -0.785817, mean_eps: 0.100000
407564/500000: episode: 492, duration: 7.012s, episode steps: 792, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.702 [0.000, 5.000], mean observation: 105.590 [87.000, 236.000], loss: 0.010116, mean_absolute_error: 0.666813, mean_q: -0.788336, mean_eps: 0.100000
408356/500000: episode: 493, duration: 6.993s, episode steps: 792, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.712 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.011620, mean_absolute_error: 0.668281, mean_q: -0.788707, mean_eps: 0.100000
409198/500000: episode: 494, duration: 7.365s, episode steps: 842, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 1.777 [0.000, 5.000], mean observation: 105.523 [87.000, 236.000], loss: 0.011835, mean_absolute_error: 0.672391, mean_q: -0.792884, mean_eps: 0.100000
410010/500000: episode: 495, duration: 7.321s, episode steps: 812, steps per second: 111, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 1.613 [0.000, 5.000], mean observation: 105.559 [87.000, 236.000], loss: 0.011869, mean_absolute_error: 0.671764, mean_q: -0.789690, mean_eps: 0.100000
410774/500000: episode: 496, duration: 6.857s, episode steps: 764, steps per second: 111, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.372 [0.000, 5.000], mean observation: 105.608 [87.000, 236.000], loss: 0.011729, mean_absolute_error: 0.673093, mean_q: -0.792699, mean_eps: 0.100000
411566/500000: episode: 497, duration: 6.962s, episode steps: 792, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.846 [0.000, 5.000], mean observation: 105.548 [87.000, 236.000], loss: 0.013258, mean_absolute_error: 0.679341, mean_q: -0.800051, mean_eps: 0.100000
412330/500000: episode: 498, duration: 6.674s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.411 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.012570, mean_absolute_error: 0.681283, mean_q: -0.800732, mean_eps: 0.100000
413094/500000: episode: 499, duration: 6.730s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.111 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.011935, mean_absolute_error: 0.681321, mean_q: -0.801414, mean_eps: 0.100000
413877/500000: episode: 500, duration: 6.788s, episode steps: 783, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.513 [0.000, 5.000], mean observation: 105.559 [87.000, 236.000], loss: 0.012267, mean_absolute_error: 0.680595, mean_q: -0.800735, mean_eps: 0.100000
414701/500000: episode: 501, duration: 7.350s, episode steps: 824, steps per second: 112, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.143 [0.000, 5.000], mean observation: 105.573 [87.000, 236.000], loss: 0.013273, mean_absolute_error: 0.685816, mean_q: -0.805882, mean_eps: 0.100000
415521/500000: episode: 502, duration: 7.127s, episode steps: 820, steps per second: 115, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 2.450 [0.000, 5.000], mean observation: 105.579 [87.000, 236.000], loss: 0.010778, mean_absolute_error: 0.676060, mean_q: -0.795758, mean_eps: 0.100000
416285/500000: episode: 503, duration: 6.683s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.236 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.012574, mean_absolute_error: 0.681727, mean_q: -0.797863, mean_eps: 0.100000
417123/500000: episode: 504, duration: 7.342s, episode steps: 838, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 3.222 [0.000, 5.000], mean observation: 105.528 [87.000, 236.000], loss: 0.013509, mean_absolute_error: 0.679993, mean_q: -0.798551, mean_eps: 0.100000
417961/500000: episode: 505, duration: 7.320s, episode steps: 838, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 4.732 [0.000, 5.000], mean observation: 105.486 [87.000, 236.000], loss: 0.011651, mean_absolute_error: 0.685496, mean_q: -0.803633, mean_eps: 0.100000
418785/500000: episode: 506, duration: 7.218s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.601 [0.000, 5.000], mean observation: 105.580 [87.000, 236.000], loss: 0.012850, mean_absolute_error: 0.689381, mean_q: -0.810834, mean_eps: 0.100000
419627/500000: episode: 507, duration: 7.410s, episode steps: 842, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 1.531 [0.000, 5.000], mean observation: 105.430 [87.000, 236.000], loss: 0.010612, mean_absolute_error: 0.675059, mean_q: -0.795236, mean_eps: 0.100000
420419/500000: episode: 508, duration: 6.917s, episode steps: 792, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.807 [0.000, 5.000], mean observation: 105.545 [87.000, 236.000], loss: 0.012597, mean_absolute_error: 0.680412, mean_q: -0.800950, mean_eps: 0.100000
421183/500000: episode: 509, duration: 6.743s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.774 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.013036, mean_absolute_error: 0.691196, mean_q: -0.813310, mean_eps: 0.100000
422008/500000: episode: 510, duration: 7.208s, episode steps: 825, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 4.105 [0.000, 5.000], mean observation: 105.547 [87.000, 236.000], loss: 0.011965, mean_absolute_error: 0.694157, mean_q: -0.816360, mean_eps: 0.100000
422846/500000: episode: 511, duration: 7.366s, episode steps: 838, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 3.780 [0.000, 5.000], mean observation: 105.509 [87.000, 236.000], loss: 0.013374, mean_absolute_error: 0.698604, mean_q: -0.820262, mean_eps: 0.100000
423638/500000: episode: 512, duration: 6.953s, episode steps: 792, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.831 [0.000, 5.000], mean observation: 105.579 [87.000, 236.000], loss: 0.010941, mean_absolute_error: 0.692848, mean_q: -0.816147, mean_eps: 0.100000
424584/500000: episode: 513, duration: 8.353s, episode steps: 946, steps per second: 113, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 2.649 [0.000, 5.000], mean observation: 105.592 [87.000, 236.000], loss: 0.012789, mean_absolute_error: 0.690894, mean_q: -0.810330, mean_eps: 0.100000
425410/500000: episode: 514, duration: 7.294s, episode steps: 826, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.946 [0.000, 5.000], mean observation: 105.549 [87.000, 236.000], loss: 0.012282, mean_absolute_error: 0.688486, mean_q: -0.808445, mean_eps: 0.100000
426234/500000: episode: 515, duration: 7.216s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.021 [0.000, 5.000], mean observation: 105.573 [87.000, 236.000], loss: 0.013494, mean_absolute_error: 0.690364, mean_q: -0.811732, mean_eps: 0.100000
427095/500000: episode: 516, duration: 7.552s, episode steps: 861, steps per second: 114, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 2.341 [0.000, 5.000], mean observation: 105.385 [87.000, 236.000], loss: 0.012090, mean_absolute_error: 0.691897, mean_q: -0.813199, mean_eps: 0.100000
427999/500000: episode: 517, duration: 7.910s, episode steps: 904, steps per second: 114, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 1.928 [0.000, 5.000], mean observation: 105.583 [87.000, 236.000], loss: 0.010911, mean_absolute_error: 0.687410, mean_q: -0.809779, mean_eps: 0.100000
428822/500000: episode: 518, duration: 7.245s, episode steps: 823, steps per second: 114, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 1.905 [0.000, 5.000], mean observation: 105.575 [87.000, 236.000], loss: 0.011874, mean_absolute_error: 0.693562, mean_q: -0.817203, mean_eps: 0.100000
429614/500000: episode: 519, duration: 6.899s, episode steps: 792, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.378 [0.000, 5.000], mean observation: 105.605 [87.000, 236.000], loss: 0.011445, mean_absolute_error: 0.689878, mean_q: -0.813280, mean_eps: 0.100000
430438/500000: episode: 520, duration: 7.424s, episode steps: 824, steps per second: 111, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.024 [0.000, 5.000], mean observation: 105.597 [87.000, 236.000], loss: 0.010875, mean_absolute_error: 0.694511, mean_q: -0.821384, mean_eps: 0.100000
431262/500000: episode: 521, duration: 7.389s, episode steps: 824, steps per second: 112, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.284 [0.000, 5.000], mean observation: 105.610 [87.000, 236.000], loss: 0.011936, mean_absolute_error: 0.702358, mean_q: -0.826133, mean_eps: 0.100000
432086/500000: episode: 522, duration: 7.258s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.523 [0.000, 5.000], mean observation: 105.623 [87.000, 236.000], loss: 0.011688, mean_absolute_error: 0.704648, mean_q: -0.827854, mean_eps: 0.100000
432910/500000: episode: 523, duration: 7.243s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.347 [0.000, 5.000], mean observation: 105.633 [87.000, 236.000], loss: 0.011885, mean_absolute_error: 0.706278, mean_q: -0.831851, mean_eps: 0.100000
433734/500000: episode: 524, duration: 7.431s, episode steps: 824, steps per second: 111, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.737 [0.000, 5.000], mean observation: 105.618 [87.000, 236.000], loss: 0.011874, mean_absolute_error: 0.704155, mean_q: -0.825179, mean_eps: 0.100000
434799/500000: episode: 525, duration: 9.402s, episode steps: 1065, steps per second: 113, episode reward: -21.000, mean reward: -0.020 [-1.000, 0.000], mean action: 1.556 [0.000, 5.000], mean observation: 105.602 [87.000, 236.000], loss: 0.012679, mean_absolute_error: 0.708379, mean_q: -0.833324, mean_eps: 0.100000
435683/500000: episode: 526, duration: 7.839s, episode steps: 884, steps per second: 113, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 1.698 [0.000, 5.000], mean observation: 105.593 [87.000, 236.000], loss: 0.014529, mean_absolute_error: 0.710861, mean_q: -0.836391, mean_eps: 0.100000
436447/500000: episode: 527, duration: 6.736s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 4.465 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.012228, mean_absolute_error: 0.709114, mean_q: -0.833587, mean_eps: 0.100000
437267/500000: episode: 528, duration: 7.193s, episode steps: 820, steps per second: 114, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 1.512 [0.000, 5.000], mean observation: 105.596 [87.000, 236.000], loss: 0.012527, mean_absolute_error: 0.707709, mean_q: -0.833148, mean_eps: 0.100000
438119/500000: episode: 529, duration: 7.510s, episode steps: 852, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.059 [0.000, 5.000], mean observation: 105.608 [87.000, 236.000], loss: 0.010031, mean_absolute_error: 0.702912, mean_q: -0.828538, mean_eps: 0.100000
438883/500000: episode: 530, duration: 6.662s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.809 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.012193, mean_absolute_error: 0.706625, mean_q: -0.832368, mean_eps: 0.100000
439727/500000: episode: 531, duration: 7.371s, episode steps: 844, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 1.950 [0.000, 5.000], mean observation: 105.531 [87.000, 236.000], loss: 0.011473, mean_absolute_error: 0.706195, mean_q: -0.832526, mean_eps: 0.100000
440491/500000: episode: 532, duration: 6.815s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.051 [0.000, 5.000], mean observation: 105.591 [87.000, 236.000], loss: 0.012788, mean_absolute_error: 0.709035, mean_q: -0.837001, mean_eps: 0.100000
441315/500000: episode: 533, duration: 7.186s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.039 [0.000, 5.000], mean observation: 105.579 [87.000, 236.000], loss: 0.011521, mean_absolute_error: 0.718808, mean_q: -0.845073, mean_eps: 0.100000
442079/500000: episode: 534, duration: 6.773s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.249 [0.000, 5.000], mean observation: 105.567 [87.000, 236.000], loss: 0.010829, mean_absolute_error: 0.717287, mean_q: -0.847009, mean_eps: 0.100000
442905/500000: episode: 535, duration: 7.268s, episode steps: 826, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.240 [0.000, 5.000], mean observation: 105.540 [87.000, 236.000], loss: 0.010524, mean_absolute_error: 0.717016, mean_q: -0.846405, mean_eps: 0.100000
443807/500000: episode: 536, duration: 7.837s, episode steps: 902, steps per second: 115, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 1.282 [0.000, 5.000], mean observation: 105.556 [87.000, 236.000], loss: 0.011819, mean_absolute_error: 0.717996, mean_q: -0.845749, mean_eps: 0.100000
444694/500000: episode: 537, duration: 7.897s, episode steps: 887, steps per second: 112, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 1.771 [0.000, 5.000], mean observation: 105.547 [87.000, 236.000], loss: 0.011445, mean_absolute_error: 0.718070, mean_q: -0.848019, mean_eps: 0.100000
445486/500000: episode: 538, duration: 7.296s, episode steps: 792, steps per second: 109, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.119 [0.000, 5.000], mean observation: 105.610 [87.000, 236.000], loss: 0.011891, mean_absolute_error: 0.717821, mean_q: -0.846151, mean_eps: 0.100000
446430/500000: episode: 539, duration: 8.519s, episode steps: 944, steps per second: 111, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 0.895 [0.000, 5.000], mean observation: 105.601 [87.000, 236.000], loss: 0.011544, mean_absolute_error: 0.718863, mean_q: -0.847432, mean_eps: 0.100000
447317/500000: episode: 540, duration: 7.728s, episode steps: 887, steps per second: 115, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 0.999 [0.000, 5.000], mean observation: 105.592 [87.000, 236.000], loss: 0.011864, mean_absolute_error: 0.718999, mean_q: -0.848308, mean_eps: 0.100000
448081/500000: episode: 541, duration: 6.791s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.632 [0.000, 5.000], mean observation: 105.592 [87.000, 236.000], loss: 0.012739, mean_absolute_error: 0.719502, mean_q: -0.848646, mean_eps: 0.100000
448919/500000: episode: 542, duration: 7.383s, episode steps: 838, steps per second: 113, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.316 [0.000, 5.000], mean observation: 105.393 [87.000, 236.000], loss: 0.012352, mean_absolute_error: 0.721368, mean_q: -0.848581, mean_eps: 0.100000
449683/500000: episode: 543, duration: 6.754s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.696 [0.000, 5.000], mean observation: 105.582 [87.000, 236.000], loss: 0.010993, mean_absolute_error: 0.719694, mean_q: -0.844900, mean_eps: 0.100000
450447/500000: episode: 544, duration: 6.703s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.283 [0.000, 5.000], mean observation: 105.578 [87.000, 236.000], loss: 0.010282, mean_absolute_error: 0.718296, mean_q: -0.847168, mean_eps: 0.100000
451239/500000: episode: 545, duration: 6.937s, episode steps: 792, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.903 [0.000, 5.000], mean observation: 105.576 [87.000, 236.000], loss: 0.012275, mean_absolute_error: 0.724525, mean_q: -0.855691, mean_eps: 0.100000
452003/500000: episode: 546, duration: 6.731s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.582 [0.000, 5.000], mean observation: 105.594 [87.000, 236.000], loss: 0.012320, mean_absolute_error: 0.727541, mean_q: -0.857000, mean_eps: 0.100000
452845/500000: episode: 547, duration: 7.488s, episode steps: 842, steps per second: 112, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 1.949 [0.000, 5.000], mean observation: 105.566 [87.000, 236.000], loss: 0.011466, mean_absolute_error: 0.727306, mean_q: -0.856382, mean_eps: 0.100000
453683/500000: episode: 548, duration: 7.385s, episode steps: 838, steps per second: 113, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 4.617 [0.000, 5.000], mean observation: 105.550 [87.000, 236.000], loss: 0.010220, mean_absolute_error: 0.724275, mean_q: -0.854537, mean_eps: 0.100000
454447/500000: episode: 549, duration: 6.703s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.805 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.010674, mean_absolute_error: 0.723826, mean_q: -0.854458, mean_eps: 0.100000
455271/500000: episode: 550, duration: 7.220s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.987 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.014370, mean_absolute_error: 0.731765, mean_q: -0.861309, mean_eps: 0.100000
456035/500000: episode: 551, duration: 6.743s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.065 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.012650, mean_absolute_error: 0.731980, mean_q: -0.863406, mean_eps: 0.100000
456799/500000: episode: 552, duration: 6.757s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.302 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.013425, mean_absolute_error: 0.741198, mean_q: -0.871429, mean_eps: 0.100000
457669/500000: episode: 553, duration: 7.747s, episode steps: 870, steps per second: 112, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 1.248 [0.000, 5.000], mean observation: 105.486 [87.000, 236.000], loss: 0.010705, mean_absolute_error: 0.728646, mean_q: -0.857996, mean_eps: 0.100000
458433/500000: episode: 554, duration: 6.835s, episode steps: 764, steps per second: 112, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.840 [0.000, 5.000], mean observation: 105.578 [87.000, 236.000], loss: 0.012782, mean_absolute_error: 0.732363, mean_q: -0.862287, mean_eps: 0.100000
459277/500000: episode: 555, duration: 7.395s, episode steps: 844, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 1.969 [0.000, 5.000], mean observation: 105.583 [87.000, 236.000], loss: 0.011205, mean_absolute_error: 0.733581, mean_q: -0.864132, mean_eps: 0.100000
460101/500000: episode: 556, duration: 7.235s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.095 [0.000, 5.000], mean observation: 105.559 [87.000, 236.000], loss: 0.013510, mean_absolute_error: 0.731827, mean_q: -0.861475, mean_eps: 0.100000
460925/500000: episode: 557, duration: 7.228s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.252 [0.000, 5.000], mean observation: 105.573 [87.000, 236.000], loss: 0.012449, mean_absolute_error: 0.736476, mean_q: -0.871065, mean_eps: 0.100000
461689/500000: episode: 558, duration: 6.672s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.908 [0.000, 5.000], mean observation: 105.569 [87.000, 236.000], loss: 0.011867, mean_absolute_error: 0.743934, mean_q: -0.875289, mean_eps: 0.100000
462453/500000: episode: 559, duration: 6.788s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.313 [0.000, 5.000], mean observation: 105.565 [87.000, 236.000], loss: 0.011833, mean_absolute_error: 0.745229, mean_q: -0.877351, mean_eps: 0.100000
463277/500000: episode: 560, duration: 7.259s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.973 [0.000, 5.000], mean observation: 105.547 [87.000, 236.000], loss: 0.012231, mean_absolute_error: 0.740547, mean_q: -0.869700, mean_eps: 0.100000
464319/500000: episode: 561, duration: 9.141s, episode steps: 1042, steps per second: 114, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action: 0.550 [0.000, 5.000], mean observation: 105.495 [87.000, 236.000], loss: 0.011064, mean_absolute_error: 0.737716, mean_q: -0.867718, mean_eps: 0.100000
465161/500000: episode: 562, duration: 7.408s, episode steps: 842, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 3.987 [0.000, 5.000], mean observation: 105.445 [87.000, 236.000], loss: 0.015903, mean_absolute_error: 0.745534, mean_q: -0.878724, mean_eps: 0.100000
466046/500000: episode: 563, duration: 7.752s, episode steps: 885, steps per second: 114, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 1.697 [0.000, 5.000], mean observation: 105.563 [87.000, 236.000], loss: 0.012779, mean_absolute_error: 0.744680, mean_q: -0.879884, mean_eps: 0.100000
466870/500000: episode: 564, duration: 7.253s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.808 [0.000, 5.000], mean observation: 105.585 [87.000, 236.000], loss: 0.010532, mean_absolute_error: 0.739721, mean_q: -0.871871, mean_eps: 0.100000
467708/500000: episode: 565, duration: 7.311s, episode steps: 838, steps per second: 115, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 1.576 [0.000, 5.000], mean observation: 105.507 [87.000, 236.000], loss: 0.012795, mean_absolute_error: 0.742557, mean_q: -0.875724, mean_eps: 0.100000
468472/500000: episode: 566, duration: 6.959s, episode steps: 764, steps per second: 110, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.891 [0.000, 5.000], mean observation: 105.574 [87.000, 236.000], loss: 0.012867, mean_absolute_error: 0.743910, mean_q: -0.880187, mean_eps: 0.100000
469356/500000: episode: 567, duration: 7.872s, episode steps: 884, steps per second: 112, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 0.510 [0.000, 5.000], mean observation: 105.594 [87.000, 236.000], loss: 0.011215, mean_absolute_error: 0.740046, mean_q: -0.872375, mean_eps: 0.100000
470120/500000: episode: 568, duration: 6.696s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.517 [0.000, 5.000], mean observation: 105.629 [87.000, 236.000], loss: 0.010915, mean_absolute_error: 0.737755, mean_q: -0.868192, mean_eps: 0.100000
471004/500000: episode: 569, duration: 7.825s, episode steps: 884, steps per second: 113, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.442 [0.000, 5.000], mean observation: 105.602 [87.000, 236.000], loss: 0.012300, mean_absolute_error: 0.748654, mean_q: -0.881713, mean_eps: 0.100000
471768/500000: episode: 570, duration: 6.711s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.666 [0.000, 5.000], mean observation: 105.623 [87.000, 236.000], loss: 0.013256, mean_absolute_error: 0.753253, mean_q: -0.885150, mean_eps: 0.100000
472772/500000: episode: 571, duration: 8.912s, episode steps: 1004, steps per second: 113, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean action: 0.273 [0.000, 5.000], mean observation: 105.642 [87.000, 236.000], loss: 0.011178, mean_absolute_error: 0.748104, mean_q: -0.880543, mean_eps: 0.100000
473536/500000: episode: 572, duration: 6.642s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.164 [0.000, 5.000], mean observation: 105.590 [87.000, 236.000], loss: 0.013489, mean_absolute_error: 0.751121, mean_q: -0.885182, mean_eps: 0.100000
474300/500000: episode: 573, duration: 6.728s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.546 [0.000, 5.000], mean observation: 105.590 [87.000, 236.000], loss: 0.012238, mean_absolute_error: 0.751271, mean_q: -0.882600, mean_eps: 0.100000
475064/500000: episode: 574, duration: 6.733s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.618 [0.000, 5.000], mean observation: 105.569 [87.000, 236.000], loss: 0.010433, mean_absolute_error: 0.750429, mean_q: -0.888667, mean_eps: 0.100000
475847/500000: episode: 575, duration: 6.819s, episode steps: 783, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.633 [0.000, 5.000], mean observation: 105.561 [87.000, 236.000], loss: 0.012523, mean_absolute_error: 0.752115, mean_q: -0.888457, mean_eps: 0.100000
476671/500000: episode: 576, duration: 7.172s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.540 [0.000, 5.000], mean observation: 105.586 [87.000, 236.000], loss: 0.011375, mean_absolute_error: 0.751751, mean_q: -0.886657, mean_eps: 0.100000
477463/500000: episode: 577, duration: 7.016s, episode steps: 792, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.922 [0.000, 5.000], mean observation: 105.595 [87.000, 236.000], loss: 0.012036, mean_absolute_error: 0.750008, mean_q: -0.885444, mean_eps: 0.100000
478227/500000: episode: 578, duration: 6.734s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.855 [0.000, 5.000], mean observation: 105.599 [87.000, 236.000], loss: 0.011527, mean_absolute_error: 0.752778, mean_q: -0.886524, mean_eps: 0.100000
479051/500000: episode: 579, duration: 7.180s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 4.076 [0.000, 5.000], mean observation: 105.526 [87.000, 236.000], loss: 0.012867, mean_absolute_error: 0.753726, mean_q: -0.886326, mean_eps: 0.100000
479889/500000: episode: 580, duration: 7.344s, episode steps: 838, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 4.521 [0.000, 5.000], mean observation: 105.384 [87.000, 236.000], loss: 0.010103, mean_absolute_error: 0.747877, mean_q: -0.882713, mean_eps: 0.100000
480653/500000: episode: 581, duration: 6.929s, episode steps: 764, steps per second: 110, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.525 [0.000, 5.000], mean observation: 105.583 [87.000, 236.000], loss: 0.011225, mean_absolute_error: 0.758814, mean_q: -0.898347, mean_eps: 0.100000
481447/500000: episode: 582, duration: 7.056s, episode steps: 794, steps per second: 113, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 2.571 [0.000, 5.000], mean observation: 105.541 [87.000, 236.000], loss: 0.012099, mean_absolute_error: 0.766777, mean_q: -0.905069, mean_eps: 0.100000
482271/500000: episode: 583, duration: 7.184s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.624 [0.000, 5.000], mean observation: 105.577 [87.000, 236.000], loss: 0.012594, mean_absolute_error: 0.772463, mean_q: -0.907662, mean_eps: 0.100000
483220/500000: episode: 584, duration: 8.312s, episode steps: 949, steps per second: 114, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 2.338 [0.000, 5.000], mean observation: 105.621 [87.000, 236.000], loss: 0.011176, mean_absolute_error: 0.768284, mean_q: -0.904616, mean_eps: 0.100000
483984/500000: episode: 585, duration: 6.658s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.260 [0.000, 5.000], mean observation: 105.625 [87.000, 236.000], loss: 0.013463, mean_absolute_error: 0.772617, mean_q: -0.908231, mean_eps: 0.100000
484748/500000: episode: 586, duration: 6.769s, episode steps: 764, steps per second: 113, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 0.836 [0.000, 5.000], mean observation: 105.599 [87.000, 236.000], loss: 0.012618, mean_absolute_error: 0.769484, mean_q: -0.907112, mean_eps: 0.100000
485636/500000: episode: 587, duration: 7.785s, episode steps: 888, steps per second: 114, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 3.051 [0.000, 5.000], mean observation: 105.541 [87.000, 236.000], loss: 0.012919, mean_absolute_error: 0.771583, mean_q: -0.909462, mean_eps: 0.100000
486474/500000: episode: 588, duration: 7.346s, episode steps: 838, steps per second: 114, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 4.684 [0.000, 5.000], mean observation: 105.549 [87.000, 236.000], loss: 0.013045, mean_absolute_error: 0.771415, mean_q: -0.906261, mean_eps: 0.100000
487238/500000: episode: 589, duration: 6.659s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.217 [0.000, 5.000], mean observation: 105.567 [87.000, 236.000], loss: 0.011300, mean_absolute_error: 0.767024, mean_q: -0.905572, mean_eps: 0.100000
488062/500000: episode: 590, duration: 7.232s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.643 [0.000, 5.000], mean observation: 105.566 [87.000, 236.000], loss: 0.013387, mean_absolute_error: 0.767763, mean_q: -0.906452, mean_eps: 0.100000
488886/500000: episode: 591, duration: 7.192s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 1.524 [0.000, 5.000], mean observation: 105.587 [87.000, 236.000], loss: 0.011801, mean_absolute_error: 0.766954, mean_q: -0.906340, mean_eps: 0.100000
489710/500000: episode: 592, duration: 7.247s, episode steps: 824, steps per second: 114, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 0.788 [0.000, 5.000], mean observation: 105.602 [87.000, 236.000], loss: 0.014167, mean_absolute_error: 0.773519, mean_q: -0.909761, mean_eps: 0.100000
490474/500000: episode: 593, duration: 6.697s, episode steps: 764, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.920 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.011294, mean_absolute_error: 0.771694, mean_q: -0.908520, mean_eps: 0.100000
491238/500000: episode: 594, duration: 6.659s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.541 [0.000, 5.000], mean observation: 105.567 [87.000, 236.000], loss: 0.012318, mean_absolute_error: 0.777658, mean_q: -0.915590, mean_eps: 0.100000
492030/500000: episode: 595, duration: 6.922s, episode steps: 792, steps per second: 114, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.032 [0.000, 5.000], mean observation: 105.549 [87.000, 236.000], loss: 0.013017, mean_absolute_error: 0.785217, mean_q: -0.923974, mean_eps: 0.100000
492856/500000: episode: 596, duration: 7.320s, episode steps: 826, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.317 [0.000, 5.000], mean observation: 105.548 [87.000, 236.000], loss: 0.010637, mean_absolute_error: 0.779532, mean_q: -0.917721, mean_eps: 0.100000
493680/500000: episode: 597, duration: 7.298s, episode steps: 824, steps per second: 113, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.853 [0.000, 5.000], mean observation: 105.548 [87.000, 236.000], loss: 0.011561, mean_absolute_error: 0.779826, mean_q: -0.920013, mean_eps: 0.100000
494504/500000: episode: 598, duration: 7.162s, episode steps: 824, steps per second: 115, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 3.661 [0.000, 5.000], mean observation: 105.539 [87.000, 236.000], loss: 0.012598, mean_absolute_error: 0.785546, mean_q: -0.925611, mean_eps: 0.100000
495268/500000: episode: 599, duration: 6.660s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 3.556 [0.000, 5.000], mean observation: 105.568 [87.000, 236.000], loss: 0.011372, mean_absolute_error: 0.784272, mean_q: -0.922880, mean_eps: 0.100000
496380/500000: episode: 600, duration: 9.800s, episode steps: 1112, steps per second: 113, episode reward: -20.000, mean reward: -0.018 [-1.000, 1.000], mean action: 1.047 [0.000, 5.000], mean observation: 105.480 [87.000, 236.000], loss: 0.011880, mean_absolute_error: 0.780382, mean_q: -0.920729, mean_eps: 0.100000
497222/500000: episode: 601, duration: 7.325s, episode steps: 842, steps per second: 115, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 1.152 [0.000, 5.000], mean observation: 105.520 [87.000, 236.000], loss: 0.012879, mean_absolute_error: 0.780437, mean_q: -0.919194, mean_eps: 0.100000
498014/500000: episode: 602, duration: 6.894s, episode steps: 792, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.129 [0.000, 5.000], mean observation: 105.598 [87.000, 236.000], loss: 0.011855, mean_absolute_error: 0.781690, mean_q: -0.916292, mean_eps: 0.100000
498834/500000: episode: 603, duration: 7.240s, episode steps: 820, steps per second: 113, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 0.749 [0.000, 5.000], mean observation: 105.598 [87.000, 236.000], loss: 0.012273, mean_absolute_error: 0.781867, mean_q: -0.917874, mean_eps: 0.100000
499598/500000: episode: 604, duration: 6.621s, episode steps: 764, steps per second: 115, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 1.179 [0.000, 5.000], mean observation: 105.619 [87.000, 236.000], loss: 0.010924, mean_absolute_error: 0.778067, mean_q: -0.915679, mean_eps: 0.100000
done, took 4351.065 seconds
Pesos guardados en: 202001081204_dqn_PongDeterministic-v0_weights.h5f
```





```
Testing for 10 episodes ...
Episode 1: reward: -21.000, steps: 764
Episode 2: reward: -21.000, steps: 764
Episode 3: reward: -21.000, steps: 764
Episode 4: reward: -21.000, steps: 764
Episode 5: reward: -21.000, steps: 764
Episode 6: reward: -21.000, steps: 764
Episode 7: reward: -21.000, steps: 764
Episode 8: reward: -21.000, steps: 764
Episode 9: reward: -21.000, steps: 764
Episode 10: reward: -21.000, steps: 764
mean reward: -21.0
mean steps: 764
(-21.0, 764)
```