Experiment 2

March 12, 2018

Simulation Details

Considered $K = 3$, $T = 1001$, $N = 500$. Report statistics at $t = 1000$
The Bandit priors that were considered:

- Uniform: Draw the mean rewards for the arms from [0.25, 0.75]

- "HeavyTail": We took the mean rewards to be randomly drawn from Beta($\alpha = 0.6, \beta = 0.6$). With this distribution it was likely to have arms that were at the extremes (close to 1 and close to 0) but also some of the arms with intermediate value means.

- Needle-in-haystack

    1. High - 2 arms with mean 0.50, 1 arm with mean 0.70 (+ 0.20)

Algorithms considered:

1. ThompsonSampling with priors of $Beta(1, 1)$ for every arm.

2. DynamicGreedy with priors of $Beta(1, 1)$ for every arm

3. Bayesian Dynamic ϵ-greedy with priors of $Beta(1, 1)$ for every arm and $\epsilon = 0.05$

Agent Algorithms considered:

1. HardMax

2. HardMaxWithRandom

3. SoftMax ($\alpha = 10$)

Memory Sizes

1. 100

Simulation Procedure

1: **for** Each prior p **do**
2:      Generate true distribution from p (except for needle-in-haystack, just use p itself)
3:      Generate $T \times K$ realizations for the arms
4:      **for** Each agent algorithm *agentalg* **do**
5:          **for** Each principal algorithm pair *principalalg1*, *principalalg2* **do**
6:              **for** N simulations **do**
7:                  Give the agents 5 observations from each principal
8:                  Give principal 2 200 free observations (the agents also get these observations)
9:                  Run simulation for T periods
10:              **end for**
11:          **end for**
12:      **end for**
13: **end for**

Results

All results are reported for memory size = 100.

The rows represent Principal 1 and the columns represent Principal 2. Thus, the cell (1, 2) represents principal 1 playing Thompson Sampling and principal 2 playing DynamicEpsilonGreedy. In this experiment remember that principal 2 gets 200 free observations.

Within each cell, the following data are presented:

1. The first row displays the sample mean market share for principal 1 as well as a 95% confidence band of the mean.

2. The second row displays the sample variance of the market share and in parentheses are 95 % confidence bands for the variance.

3. The third row displays the % of simulations that resulted in "extreme" market shares. Extreme market shares are defined as being simulations were one of the principals (either principal 1 or 2) ended up with 90% or more of the market.

Results for HardMax t = 1000 Needle In Haystack High

	ThompsonSampling	DynamicEpsilonGreedy	DynamicGreedy
ThompsonSampling	**0.11** +/- 0.023 0.066 (0.059, 0.075) Extreme Shares: 88 %	**0.15** +/- 0.028 0.1 (0.092, 0.12) Extreme Shares: 90 %	**0.23** +/- 0.034 0.15 (0.13, 0.17) Extreme Shares: 87 %
DynamicEpsilonGreedy	**0.1** +/- 0.023 0.066 (0.059, 0.075) Extreme Shares: 89 %	**0.17** +/- 0.029 0.11 (0.1, 0.13) Extreme Shares: 85 %	**0.27** +/- 0.036 0.17 (0.15, 0.19) Extreme Shares: 86 %
DynamicGreedy	**0.11** +/- 0.022 0.063 (0.056, 0.072) Extreme Shares: 85 %	**0.16** +/- 0.029 0.11 (0.096, 0.12) Extreme Shares: 87 %	**0.23** +/- 0.033 0.14 (0.13, 0.16) Extreme Shares: 85 %

Results for HardMaxWithRandom t = 1000 Needle In Haystack High

	ThompsonSampling	DynamicEpsilonGreedy	DynamicGreedy
ThompsonSampling	**0.17** +/- 0.022 0.064 (0.057, 0.073) Extreme Shares: 72 %	**0.24** +/- 0.028 0.1 (0.091, 0.12) Extreme Shares: 69 %	**0.35** +/- 0.034 0.15 (0.14, 0.17) Extreme Shares: 71 %
DynamicEpsilonGreedy	**0.24** +/- 0.026 0.091 (0.08, 0.1) Extreme Shares: 63 %	**0.26** +/- 0.029 0.11 (0.098, 0.13) Extreme Shares: 67 %	**0.37** +/- 0.033 0.14 (0.12, 0.16) Extreme Shares: 62 %
DynamicGreedy	**0.21** +/- 0.025 0.079 (0.07, 0.09) Extreme Shares: 65 %	**0.28** +/- 0.03 0.11 (0.1, 0.13) Extreme Shares: 66 %	**0.37** +/- 0.033 0.14 (0.13, 0.16) Extreme Shares: 64 %

Results for SoftMax t = 1000 Needle In Haystack High

	ThompsonSampling	DynamicEpsilonGreedy	DynamicGreedy
ThompsonSampling	**0.43** +/- 0.008 0.0082 (0.0073, 0.0093) Extreme Shares: 0 %	**0.48** +/- 0.012 0.019 (0.016, 0.021) Extreme Shares: 0 %	**0.51** +/- 0.016 0.031 (0.028, 0.036) Extreme Shares: 0 %
DynamicEpsilonGreedy	**0.41** +/- 0.011 0.017 (0.015, 0.019) Extreme Shares: 0 %	**0.44** +/- 0.014 0.026 (0.023, 0.03) Extreme Shares: 0 %	**0.51** +/- 0.017 0.039 (0.034, 0.044) Extreme Shares: 0 %
DynamicGreedy	**0.4** +/- 0.012 0.02 (0.018, 0.023) Extreme Shares: 0 %	**0.44** +/- 0.015 0.031 (0.027, 0.035) Extreme Shares: 0.2 %	**0.47** +/- 0.018 0.04 (0.035, 0.045) Extreme Shares: 0 %

Results for HardMax t = 1000 Heavy Tail

	ThompsonSampling	**DynamicEpsilonGreedy**	**DynamicGreedy**
ThompsonSampling	**0.052** +/- 0.016 0.034 (0.03, 0.039) Extreme Shares: 93 %	**0.09** +/- 0.023 0.066 (0.058, 0.075) Extreme Shares: 93 %	**0.11** +/- 0.025 0.079 (0.07, 0.089) Extreme Shares: 92 %
DynamicEpsilonGreedy	**0.082** +/- 0.019 0.045 (0.04, 0.051) Extreme Shares: 86 %	**0.15** +/- 0.026 0.085 (0.075, 0.096) Extreme Shares: 81 %	**0.16** +/- 0.026 0.086 (0.076, 0.098) Extreme Shares: 78 %
DynamicGreedy	**0.14** +/- 0.026 0.085 (0.076, 0.097) Extreme Shares: 82 %	**0.21** +/- 0.031 0.12 (0.11, 0.14) Extreme Shares: 81 %	**0.17** +/- 0.027 0.093 (0.083, 0.11) Extreme Shares: 78 %

Results for HardMaxWithRandom t = 1000 Heavy Tail

	ThompsonSampling	**DynamicEpsilonGreedy**	**DynamicGreedy**
ThompsonSampling	**0.11** +/- 0.017 0.04 (0.035, 0.045) Extreme Shares: 80 %	**0.19** +/- 0.024 0.077 (0.069, 0.088) Extreme Shares: 71 %	**0.22** +/- 0.028 0.1 (0.089, 0.11) Extreme Shares: 75 %
DynamicEpsilonGreedy	**0.17** +/- 0.022 0.06 (0.054, 0.069) Extreme Shares: 67 %	**0.25** +/- 0.027 0.092 (0.082, 0.1) Extreme Shares: 62 %	**0.24** +/- 0.027 0.091 (0.081, 0.1) Extreme Shares: 63 %
DynamicGreedy	**0.21** +/- 0.025 0.08 (0.071, 0.09) Extreme Shares: 65 %	**0.31** +/- 0.029 0.11 (0.098, 0.13) Extreme Shares: 56 %	**0.29** +/- 0.029 0.11 (0.094, 0.12) Extreme Shares: 59 %

Results for SoftMax t = 1000 Heavy Tail

	ThompsonSampling	**DynamicEpsilonGreedy**	**DynamicGreedy**
ThompsonSampling	**0.44** +/- 0.0063 0.0052 (0.0046, 0.0059) Extreme Shares: 0 %	**0.49** +/- 0.0075 0.0072 (0.0064, 0.0082) Extreme Shares: 0 %	**0.48** +/- 0.011 0.017 (0.015, 0.019) Extreme Shares: 0 %
DynamicEpsilonGreedy	**0.42** +/- 0.008 0.0083 (0.0073, 0.0094) Extreme Shares: 0 %	**0.45** +/- 0.0093 0.011 (0.01, 0.013) Extreme Shares: 0 %	**0.47** +/- 0.013 0.02 (0.018, 0.023) Extreme Shares: 0.4 %
DynamicGreedy	**0.45** +/- 0.0098 0.012 (0.011, 0.014) Extreme Shares: 0 %	**0.47** +/- 0.012 0.017 (0.015, 0.019) Extreme Shares: 0.2 %	**0.48** +/- 0.011 0.016 (0.014, 0.018) Extreme Shares: 0 %

Results for HardMax t = 1000 Uniform

	ThompsonSampling	**DynamicEpsilonGreedy**	**DynamicGreedy**
ThompsonSampling	**0.13** +/- 0.023 0.071 (0.063, 0.08) Extreme Shares: 81 %	**0.17** +/- 0.029 0.11 (0.096, 0.12) Extreme Shares: 84 %	**0.21** +/- 0.032 0.13 (0.11, 0.15) Extreme Shares: 82 %
DynamicEpsilonGreedy	**0.15** +/- 0.025 0.083 (0.074, 0.095) Extreme Shares: 79 %	**0.16** +/- 0.027 0.095 (0.084, 0.11) Extreme Shares: 82 %	**0.22** +/- 0.031 0.13 (0.11, 0.15) Extreme Shares: 81 %
DynamicGreedy	**0.15** +/- 0.026 0.087 (0.077, 0.099) Extreme Shares: 80 %	**0.18** +/- 0.029 0.11 (0.094, 0.12) Extreme Shares: 80 %	**0.24** +/- 0.032 0.13 (0.12, 0.15) Extreme Shares: 74 %

Results for HardMaxWithRandom t = 1000 Uniform

	ThompsonSampling	**DynamicEpsilonGreedy**	**DynamicGreedy**
ThompsonSampling	**0.24** +/- 0.025 0.082 (0.073, 0.094) Extreme Shares: 57 %	**0.27** +/- 0.028 0.1 (0.09, 0.12) Extreme Shares: 58 %	**0.32** +/- 0.03 0.12 (0.11, 0.13) Extreme Shares: 57 %
DynamicEpsilonGreedy	**0.27** +/- 0.026 0.089 (0.079, 0.1) Extreme Shares: 53 %	**0.28** +/- 0.026 0.088 (0.078, 0.1) Extreme Shares: 50 %	**0.35** +/- 0.03 0.12 (0.11, 0.14) Extreme Shares: 53 %
DynamicGreedy	**0.28** +/- 0.027 0.091 (0.081, 0.1) Extreme Shares: 52 %	**0.31** +/- 0.028 0.1 (0.091, 0.12) Extreme Shares: 51 %	**0.34** +/- 0.03 0.12 (0.1, 0.13) Extreme Shares: 55 %

Results for SoftMax t = 1000 Uniform

	ThompsonSampling	**DynamicEpsilonGreedy**	**DynamicGreedy**
ThompsonSampling	**0.45** +/- 0.0081 0.0084 (0.0075, 0.0096) Extreme Shares: 0 %	**0.47** +/- 0.0096 0.012 (0.011, 0.014) Extreme Shares: 0 %	**0.5** +/- 0.012 0.019 (0.017, 0.022) Extreme Shares: 0 %
DynamicEpsilonGreedy	**0.45** +/- 0.0095 0.012 (0.01, 0.013) Extreme Shares: 0 %	**0.46** +/- 0.01 0.014 (0.013, 0.016) Extreme Shares: 0 %	**0.49** +/- 0.014 0.024 (0.021, 0.027) Extreme Shares: 0 %
DynamicGreedy	**0.44** +/- 0.011 0.016 (0.014, 0.018) Extreme Shares: 0 %	**0.47** +/- 0.012 0.018 (0.016, 0.02) Extreme Shares: 0.2 %	**0.47** +/- 0.013 0.023 (0.02, 0.026) Extreme Shares: 0 %

Comments: comment