Corrigendum:
Mutual Observability and the Convergence of Actions in a Multi-Person Two-Armed Bandit Model

Masaki Aoyagi*
Osaka University

February 2, 2011

The proof of Lemma 1 in the paper claims in equation (a2) that $\hat{\alpha}^m$ converges μ_*-almost surely to 1 by the consistency of Bayes estimates. However, conditional on F_i, there is no guarantee that player i's trials on arm X are Bernoulli trials since the random outcomes on the arm may not be independent conditional on F_i. In this sense, (a2) is unwarranted. The proof below fixes this problem.

Proof of Lemma 1 Fix player i, and value $x = r_k$, and write α^t for $\alpha^t_\sigma(i,k)$: i's period t posterior belief that the value of arm X is r_k. Let q^t_{ix} be the outcome on arm X by player i in period t. Player i will observe q^t_{ix} in period t if he indeed plays arm X in that period. He does not observe q^t_{ix} if he chooses to play arm Y in period t. Let $\bar{z}^0_{ix} = s_i$, $\bar{z}^t_{ix} = (s_i, q^1_{ix}, \dots, q^t_{ix})$ and $\bar{z}^t = (\bar{z}^t_{1x}, \bar{z}^t_{1y}, \bar{z}^t_{2x}, \bar{z}^t_{2y})$ $(t \in \boldsymbol{N} \cup \{\infty\})$. Denote by μ the distribution of the random variables x, y, and $\bar{z}^\infty$. Note that μ is independent of the particular strategy profile σ. For each $m \in \boldsymbol{N}$, let $K_m \in \boldsymbol{N} \cup \{\infty\}$ be the random period in which player i plays arm X for the mth time. Set $K_m = \infty$ if arm X is played less than m times. Since each K_m is such that $\{K_m = t\}$ is measurable with respect to $\bar{z}^{t-1}$ for $t = 1, 2, \dots, \infty$, it is a stopping time relative to $\sigma(z^{t-1})$. For each $m \in \boldsymbol{N}$, let $q^{K_m}_{ix}$ be the outcome on arm X by player i in period K_m if $K_m < \infty$. Let $q^m_{ix} = -1$ if $K_m = \infty$. Define now z^m_{ix} $(m \in \boldsymbol{Z})$ by $z^0_{ix} = s_i$ and $z^m_{ix} = (s_i, q^{K_1}_{ix}, \dots, q^{K_m}_{ix})$ $(m \in \boldsymbol{N})$. Namely, z^m_{ix} is part of player i's private history recording only the outcomes of trials on arm X and the private signal

*ISER, Osaka University, 6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan.

s_i. Define $\hat{\alpha}^m = \mu(x = r_k \mid z_{ix}^m)$ to be i's period K_m posterior belief that $x = r_k$ given z_{ix}^m as the only information. Since K_m and $q_{ix}^{K_m}$ are independent conditional on (x, y) (Cho and Teicher [1, Corollary 5.3.1]), it follows that if $K_m < \infty$ for every $m < \infty$, $q_{ix}^{K_1}, q_{ix}^{K_2}, \ldots, q_{ix}^{K_m}, \ldots$ consititutes Bernoulli trials on arm X. Since $F_i = \bigcap_{m \in \boldsymbol{N}} \{K_m < \infty\}$, we have

$$\mu_\sigma(\lim_{m\to\infty} \hat{\alpha}^m = 1, F_i \mid x = r_k) = 1$$

by the consistency of Bayes estimates.[1] Since $\mu_\sigma(x = r_k, F_i) > 0$ by assumption, the above implies that

$$\mu_\sigma(\lim_{m\to\infty} \hat{\alpha}^m = 1 \mid x = r_k, F_i) = 1. \tag{1}$$

We now turn to the proof of $\mu_\sigma(\lim_{t\to\infty} \alpha^t = 1 \mid x = r_k, F_i) = 1$. Write $\alpha^\infty = \mu_\sigma(x = r_k \mid h^\infty, z_i^\infty)$ and $\hat{\alpha}^\infty = \mu(x = r_k \mid z_{ix}^\infty)$. We then have $\lim_{t\to\infty} \alpha^t = \alpha^\infty$ and $\lim_{m\to\infty} \hat{\alpha}^m = \hat{\alpha}^\infty$ both μ_σ-almost surely by the martingale convergence theorem. Let E_σ be the expectation operator corresponding to μ_σ. Since $\sigma(z_{ix}^\infty) \subset \sigma(h^\infty, z_i^\infty)$,

$$\hat{\alpha}^\infty = E_\sigma[1_{\{x=r_k\}} \mid z_{ix}^\infty] = E_\sigma[E_\sigma[1_{\{x=r_k\}} \mid h^\infty, z_i^\infty] \mid z_{ix}^\infty] = E_\sigma[\alpha^\infty \mid z_{ix}^\infty] \quad \mu_\sigma\text{-a.s.} \tag{2}$$

by the law of iterated expectations. Let $G = \{\hat{\alpha}^\infty = 1\}$ and note that by our assumption and (1),

$$\mu_\sigma(G) \geq \mu_\sigma(G \mid x = r_k, F_i)\, \mu_\sigma(x = r_k, F_i) = \mu_\sigma(x = r_k, F_i) > 0. \tag{3}$$

We then have

$$1 = E_\sigma[\hat{\alpha}^\infty \mid G] = E_\sigma[E_\sigma[\alpha^\infty \mid z_{ix}^\infty] \mid G] = E_\sigma[\alpha^\infty \mid G], \tag{4}$$

where the first equality follows from the definition of G, the second from (2), and the third from the fact that the set G is $\sigma(z_{ix}^\infty)$-measurable. Since $\alpha^\infty \leq 1$, (3) and (4) show that $\mu_\sigma(\alpha^\infty = 1 \mid G) = 1$ must be true. This and (1) imply the desired conclusion $\mu_\sigma(\alpha^\infty = 1 \mid x = r_k, F_i) = 1$. ■

References

[1] Y. S. Chow, H. Teicher, Probability Theory, Second Edition, Springer-Verlag, New York, 1988.

[2] M. DeGroot, Optimal Statistical Decisions, McGraw-Hill, New York, 1970.

[1] See, for example, DeGroot [2, Sec. 10.5].