SUPPLEMENT TO "STATISTICAL INFERENCE IN GAMES"

YUVAL SALANT
Kellogg School of Management, Northwestern University

JOSH CHERRY
Pinterest

THIS SUPPLEMENT EXTENDS the definitions of the main text to games with more than two actions, provides an equilibrium existence result, and solves an example.

Primitives. There is a unit mass of players, and each of them has to decide which of $N + 1$ actions to take. The utility from action 0 is 0. The utility from action $n \in \{1, 2, \ldots, N\}$ is $u_n(\theta_n, \alpha_n) = \theta_n - f_n(\alpha_n)$, where θ_n is a player's idiosyncratic benefit of taking action n and $f_n(\alpha_n)$ is the cost incurred by a player taking action n when a proportion α_n of players take action n. Each cost function f_n is continuous with $0 \le f_n(0)$ and $f_n(1) \le 1$.

Statistical Decision Making. In order to decide which action to take, each player estimates the vector $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_N) \in \Delta^N$. The player obtains k independent observations from a Categorical distribution with the parameter $\boldsymbol{\alpha}$. The resulting sample is $(k, \mathbf{z}) = (k, (z_0, z_1, z_2, \ldots, z_N))$, where the integer $k \ge 1$ is the sample size, $\mathbf{z} \in \Delta^N$, and $z_n \in [0, 1]$ is the proportion of "successes" for action n in the sample. The player then forms an estimate according to an inference procedure and best-responds to this estimate in taking an action.

DEFINITION: An *inference procedure* $G = \{G_{k,\mathbf{z}}\}$ assigns to every sample $(k, \mathbf{z})$ a cumulative distribution function $G_{k,\mathbf{z}}$ defined over $[0, 1]^{N+1}$, called an estimate, such that:
For any action n and any two vectors $\mathbf{z}$ and $\hat{\mathbf{z}}$ with $z_n > \hat{z}_n$, the marginal estimate of action n under $G_{k,\mathbf{z}}$ strictly first-order stochastically dominates the marginal estimate of n under $G_{k,\hat{\mathbf{z}}}$.

Following are a few examples of inference procedures.

EXAMPLE S1—Bayesian Updating: A player has a non-degenerate prior on $\boldsymbol{\alpha}$ and uses Bayes' rule to update this prior based on the sample.

EXAMPLE S2—Maximum Likelihood Estimation: A player uses the maximum likelihood method to estimate the most likely vector $\boldsymbol{\alpha}$ to generate the sample, that is,

$$\boldsymbol{\alpha}_{k,\text{MLE}} = \arg\max_{\boldsymbol{\alpha}} \frac{k!}{\prod (k z_n)!} \prod \alpha_n^{k z_n}.$$

It is easy to verify that $\boldsymbol{\alpha}_{k,\text{MLE}} = (z_0, z_1, \ldots, z_N)$.

Yuval Salant: y-salant@kellogg.northwestern.edu
Josh Cherry: josh.cherry@gmail.com

EXAMPLE S3—*Dirichlet Estimation*: This inference procedure is a generalization of the Beta estimation procedure. A player has "complete ignorance" about the vector $\boldsymbol{\alpha}$. Formally, the player's "prior" is the limit of the Dirichlet distribution with parameter $\boldsymbol{\epsilon} = (\epsilon, \epsilon, \ldots, \epsilon)$ as $\epsilon \to 0$. After observing j_n "successes" for action n in a sample of size k, the player "updates" his estimate to the Dirichlet distribution with the parameter $(j_0, j_1, \ldots, j_N)$.

Equilibrium.   In a SESI, players obtain their sample from the distribution of actions based on players' statistical decision making.

DEFINITION: A *SESI* is a vector $\boldsymbol{\alpha}_{k,G} = (\alpha_{0,(k,G)}, \alpha_{1,(k,G)}, \ldots, \alpha_{N,(k,G)}) \in \Delta^N$ such that, for any $0 \le n \le N$, an $\alpha_{n,(k,G)}$ proportion of players take action n when each player obtains k independent observations from a Categorical distribution with parameter $\boldsymbol{\alpha}_{k,G}$, forms an estimate according to the inference procedure G, and best-responds to this estimate in choosing an action.

OBSERVATION S1: *A SESI exists for every inference procedure and every sample size.*

PROOF: Fix an inference procedure G and sample size k. Consider the best-response function $\boldsymbol{h}$ that assigns to any $\boldsymbol{\alpha} \in \Delta^N$ the $N + 1$ proportions $\boldsymbol{h}(\boldsymbol{\alpha})$ of players taking each of the $N + 1$ actions after obtaining k observations from a Categorical distribution with the parameter $\boldsymbol{\alpha}$ and best-responding to G.

Every player observes one of finitely many samples with a probability that is continuous in $\boldsymbol{\alpha}$. Conditional on the sample and players' tie-breaking rule, players' probability distribution over actions is constant. Thus, the best-response function $\boldsymbol{h}$ is the sum of finitely many components, each continuous in $\boldsymbol{\alpha}$, and is therefore continuous in $\boldsymbol{\alpha}$. Hence, by Brouwer's fixed point theorem, $\boldsymbol{h}$ has a fixed point.   *Q.E.D.*

Application: Choosing Which Product Variety to Produce in a Competitive Market.   This application modifies the competitive market application in the main text.

There is a unit mass of producers and each of them has to decide which of N product varieties to produce, if any. The inverse demand $P(Q_n)$ for variety n is independent of n, it decreases in Q_n, and is between 0 and 1. Production cost θ is also independent of n, and is distributed uniformly on $[0, 1]$ among producers.

Because production cost is independent of n, each producer wishes to produce the variety with the highest market price. Because price decreases in quantity and the inverse demand is identical for different varieties, the variety with the highest price is the variety with the lowest aggregate production. Thus, each producer needs to estimate $\boldsymbol{Q} = (Q_1, \ldots, Q_N)$ in order to decide which variety to produce, if any.

Fix a sample size $k \le N - 1$ and suppose producers use MLE to estimate $\boldsymbol{Q}$. Suppose also that if a producer is indifferent between producing one of several varieties, he mixes uniformly between them. Then, the following holds:

OBSERVATION S2: *There is a unique SESI. In this SESI, all varieties are produced by the same proportion of producers. There is also a unique NE. In this NE, all varieties are produced by the same proportion of producers, which is strictly smaller than in the unique SESI.*

PROOF: To solve for a SESI and show it is unique, we observe that $k \le N - 1$ implies that any sample has one or more varieties that no one produces. In a SESI, a producer

will produce one of these varieties if $\theta \leq P(0)$. Otherwise, he will not produce. Thus, $\alpha_{0,(k,\text{MLE})} = 1 - P(0)$.

If the remaining mass $P(0)$ of producers is not divided equally among varieties in a SESI, then the variety with the largest mass of producers will appear in samples more frequently than the variety with the smallest mass. This implies that more producers will choose to produce the variety with the smallest mass. Thus, the only candidate for a SESI is a distribution in which the mass $P(0)$ is divided equally among varieties. It is easy to verify that this is indeed a SESI.

In a NE, total production mass should also be divided equally among varieties. Denoting the production mass of each variety by Q_{NE}, the total production mass NQ_{NE} has to solve $NQ = P(Q)$ because producers with $\theta \geq P(Q)$ choose not to produce. Because P decreases in Q with $P(0) > 0$ and $P(1) < 1$ while NQ increases in Q, there is a unique NE. Because $Q_{\text{NE}} > 0$, $P(Q_{\text{NE}}) < P(0)$ and hence the mass producing each variety $\frac{P(Q_{\text{NE}})}{N}$ is smaller than in the unique SESI independently of the exact shape of P. *Q.E.D.*