Deterministic Calibration and Nash Equilibrium

Sham M. Kakade

Dean P. Foster
University of Pennsylvania

Deterministic Calibration and Nash Equilibrium

Abstract

We provide a natural learning process in which the joint frequency of empirical play converges into the set of convex combinations of Nash equilibria. In this process, all players rationally choose their actions using a public prediction made by a deterministic, *weakly calibrated* algorithm. Furthermore, the public predictions used in any given round play are frequently close to some Nash equilibrium of the game.

Keywords

Nash equilibria, calibration, correlated equilibria, game theory, learning

Disciplines

Applied Statistics | Behavioral Economics | Statistics and Probability | Theory and Algorithms

Deterministic Calibration and Nash Equilibrium

Sham M. Kakade and Dean P. Foster

University of Pennsylvania, Philadelphia, PA 19104

Abstract. We provide a natural learning process in which the joint frequency of empirical play converges into the set of convex combinations of Nash equilibria. In this process, all players rationally choose their actions using a public prediction made by a deterministic, *weakly calibrated* algorithm. Furthermore, the public predictions used in any given round of play are frequently close to some Nash equilibrium of the game.

1 Introduction

Perhaps the most central question for justifying any game theoretic equilibrium as a general solution concept is: can we view the equilibrium as a convergent point of a sensible learning process? Unfortunately for Nash equilibria, there are currently no learning algorithms in the literature in which play generally converges (in some sense) to a Nash equilibrium of the one shot game, short of exhaustive search — see Foster and Young [forthcoming] for perhaps the most general result in which players sensibly search through hypothesis. In contrast, there is a long list of special cases (*eg* zero sum games, 2x2 games, assumptions about the players' prior subjective beliefs) in which there exist learning algorithms that have been shown to converge (a representative but far from exhaustive list would be Robinson [1951], Milgrom and Roberts [1991], Kalai and Lehrer [1993], Fudenberg and Levine [1998], Freund and Schapire [1999]).

If we desire that the mixed strategies themselves converge to a Nash equilibrium, then a recent result by Hart and Mas-Colell [2003] shows that this is, in general, not possible under a certain class of learning rules [1]. Instead, one can examine the convergence of the joint frequency of the empirical play, which has the advantage of being an observable quantity. This has worked well in the case of a similar equilibrium concept, namely correlated equilibrium (Foster and Vohra [1997], Hart and Mas-Colell [2000]). However, for Nash equilibria, previous general results even for this weaker form of convergence are limited to some form of exhaustive search (though see Foster and Young [forthcoming]).

In this paper, we provide a learning process in which the joint frequency of empirical play converges to a Nash equilibrium, if it is unique. More generally, convergence is into the set of convex combinations of Nash equilibria (where the empirical play could jump from one Nash equilibrium to another infinitely

[1] They show that, in general, there exists no continuous time dynamics which converge to a Nash equilibrium (even if the equilibrium is unique), with the natural restriction that a players mixed strategy is updated without using the knowledge of the other players' utility functions.

often). Our learning process is the most traditional one: players make predictions of their opponents and take best responses to their predictions. Central to our learning process is the use of public predictions formed by an "accurate" (*eg* calibrated) prediction algorithm.

We now outline the main contributions of this paper.

"Almost" Deterministic Calibration Formulating sensible prediction algorithms is a notoriously difficult task in the game theoretic setting [2]. A rather minimal requirement for any prediction algorithm is that it should be *calibrated* (see Dawid [1982]). An informal explanation of calibration would go something like this. Suppose each day a weather forecaster makes some prediction, say p, of the chance that it rains the next day. Now from the subsequence of days on which the forecaster announced p, compute the empirical frequency that it actually rained the next day, and call this $\rho(p)$. Crudely speaking, calibration requires that $\rho(p)$ equal p, if the forecast p is used often.

If the weather acts adversarially, then Oakes [1985] and Dawid [1985] show that a deterministic forecasting algorithm will not be always be calibrated. However, Foster and Vohra [1998] show that calibration is almost surely guaranteed with a randomized forecasting rule, *ie* where the forecasts are chosen using private randomization and the forecasts are hidden from the weather until the weather makes its decision to rain or not. Of course, this solution makes it difficult for a weather forecaster to publicly announce a prediction.

Although stronger notions of calibration have been proposed (see Kalai et al. [1999]), here we actually consider a weaker notion [3]. Our contribution is to provide a *deterministic* algorithm that is always *weakly calibrated*. Rather than precisely defining weak calibration here, we continue to with our example to show how this deterministic algorithm can be used to obtain calibrated forecasts in the standard sense.

Assume the weather forecaster uses our deterministic algorithm and publicly announces forecasts to a number of observers interested in the weather. Say the following forecasts are made over some period of 5 days:

$$0.8606,\ 0.2387,\ 0.57513,\ 0.4005,\ 0.069632,\ \ldots$$

How can an interested observer make calibrated predictions using this announced forecast? In our setting, an observer can just *randomly round* the forecasts in order to calibrate. For example, if the observer rounds to the second digit, then on the first day, the observer will privately predict .87 with probability .06 and .86 otherwise, and, on the second day, the private predictions will be 0.24 with probability 0.87 and 0.23 otherwise. Under this scheme, the asymptotic calibration error of the observer will, almost surely, be small (and if the observer rounded to the third digit, this error would be yet even smaller).

[2] Subjective notions of probability fall prey to a host of impossibility results — crudely, Alice wants to predict Bob while Bob wants to predict Alice, which leads to a feedback loop (if Alice and Bob are both rational). See Foster and Young [2001].

[3] We use the word "weak" in the technical sense of weak convergence of measures (see Billingsley [1968]) rather than how it used by Kalai et al. [1999].

Unlike previous calibrated algorithms, this deterministic algorithm provides a meaningful forecast, which can be calibrated using only randomized rounding.

Nash Convergence The existence of a deterministic forecasting scheme leaves open the possibility that all players can rationally use some public forecast, since each player is guaranteed to form calibrated predictions (*regardless* of how the other players behave). For example, say some public forecaster provides a prediction of the *full joint distribution* of all n players. The algorithm discussed above can be generalized such that each player can use this prediction (with randomized rounding) to construct a prediction of the other players. Each player can then use their own prediction to choose a best response.

We formalize this scheme later, but point out that our (weakly) calibrated forecasting algorithm only needs to observe the history of play (and does not require any information about the players' utility functions). Furthermore, there need not be any "publicly announced" forecast provided to every player at each round — alternatively, each player could have knowledge of the deterministic forecasting algorithm and could perform the computation themselves.

Now Foster and Vohra [1997] showed that if players make predictions that satisfy the rather minimal calibration condition, then the joint frequency of the empirical play converges into the set of correlated equilibria. Hence, it is immediate that in our setting, convergence is into the set of correlated equilibria. However, we can prove the stronger condition that the joint frequency of empirical play converges into the set of convex combinations of Nash equilibria, a smaller set than that of correlated equilibria. This directly implies that the average payoff achieved by each player is at least the player's payoff under some Nash equilibrium — a stronger guarantee than achieving a (possibly smaller) correlated equilibrium payoff.

This setting deals with the coordination problem of "which Nash equilibrium to play?" in a natural manner. The setting does not arbitrarily force play to any single equilibrium and allows the possibility that players could (jointly) switch play from one Nash equilibrium to another — perhaps infinitely often. Furthermore, although play converges to the convex combinations of Nash equilibria, we have the stronger result that the public forecasts themselves are frequently close to some Nash equilibrium (*not* general combinations of them). Of course if the Nash equilibrium is unique, then the empirical play converges to it.

The convergence rate, until the empirical play is an approximate Nash equilibrium, is $O(\sqrt{T})$ (where T is the number of rounds of play), with constants that are exponential in both the number of players and actions. Hence, our setting does not lead to a polynomial time algorithm for computing an approximate Nash equilibrium (which is currently an important open problem).

2 Deterministic Calibration

We first describe the online prediction setting. There is a finite outcome space $\Omega = \{1, 2, \ldots\ |\Omega|\}$. Let X be an infinite sequence of outcomes, whose t-th element, X_t, indicates the outcome on time t. For convenience, we represent the outcome $X_t = (X_t[1], X_t[2], \ldots X_t[|\Omega|])$ as a binary vector in $\{0,1\}^{|\Omega|}$ that indicates which state at time t was realized — if the realized state was i, then the i-th component of X_t is 1 and all other components are 0. Hence, $\frac{1}{T}\sum_{t=1}^{T} X_t$ is the empirical frequency of the outcomes up to time T and is a valid probability distribution.

A forecasting method, F, is simply a function from a sequence of outcomes to a probability distribution over Ω. The forecast that F makes in time t is denoted by $f_t = F(X_1, X_2, \ldots, X_{t-1})$ (clearly, the t-th forecast must be made without knowledge of X_t). Here $f_t = (f_t[1], f_t[2], \ldots f_t[|\Omega|])$, where the ith component is the forecasted probability that state i will be realized in time t.

2.1 Weak Calibration

We now define a quantity to determine if F is calibrated with respect to some probability distribution p. Define $I_{p,\epsilon}(f)$ to be a "test" function indicating if the forecast f is ϵ-close to p, *ie*

$$I_{p,\epsilon}(f) = \begin{cases} 1 & \text{if } |f - p| \le \epsilon \\ 0 & \text{else} \end{cases}$$

where $|f|$ denotes the l_1 norm, *ie* $|f| = \sum_{k \in \Omega} \big|f[k]\big|$. We define the calibration error μ_T of F with respect to $I_{p,\epsilon}$ as:

$$\mu_T(I_{p,\epsilon}, X, F) = \frac{1}{T}\sum_{t=1}^{T} I_{p,\epsilon}(f_t)(X_t - f_t)$$

Note that $X_t - f_t$ is the immediate error (which is a vector) and the above error μ_T measures this instantaneous error on those times when the forecast was ϵ-close to p.

We say that F is *calibrated* if for all sequences X and all test functions $I_{p,\epsilon}$, the calibration error tends to 0, *ie*

$$\mu_T(I_{p,\epsilon}, X, F) \to 0$$

as T tends to infinity. As discussed in the Introduction, there exist no deterministic rules F that are calibrated (Dawid [1985], Oakes [1985]). However, Foster and Vohra [1998] show that there exist randomized forecasting rules F (*ie* F is a randomized function) which are calibrated. Namely, there exists a randomized F such that for all sequences X and for all test functions $I_{p,\epsilon}$, the error $\mu_T(I_{p,\epsilon}, X, F) \to 0$ as T tends to infinity, with probability 1 (where the probability is taken with respect to the randomization used by the forecasting scheme).

We now generalize this definition of the calibration error by defining it with respect to arbitrary test functions w, where a *test function* is defined as a mapping from probability distributions into the interval $[0,1]$. We define the calibration error μ_T of F with respect to the test function w as:

$$\mu_T(w, X, F) = \frac{1}{T}\sum_{t=1}^{T} w(f_t)(X_t - f_t)$$

This is consistent with the previous definition if we set $w{=}I_{p,\epsilon}$.

Let W be the set of all test functions which are Lipschitz continuous functions [4]. We say that F is *weakly calibrated* if for all sequences X and all $w \in W$,

$$\mu_T(w, X, F) \to 0$$

as T tends to infinity. Also, we say that F is *uniformly, weakly calibrated* if for all $w \in W$,

$$\sup_X \mu_T(w, X, F) \to 0$$

as T tends to infinity. The latter condition is strictly stronger. Our first main result follows.

Theorem 1. *(Deterministic Calibration) There exists a deterministic forecasting rule which is uniformly, weakly calibrated.*

The proof of this theorem is constructive and is presented in section 4.

2.2 Randomized Rounding for Standard Calibration

We now show how to achieve calibration in the standard sense (with respect to the indicator functions $I_{p,\epsilon}$), using a deterministic weakly calibrated algorithm along with some randomized rounding. Essentially, the algorithm rounds any forecast to some element in a finite set, V, of forecasts. In the example in the Introduction, the set V was the set of probability distributions which are specified up to the second digit of precision.

Let Δ be the simplex in which the forecasts live ($\Delta \subset \mathcal{R}^{|\Omega|}$). Consider some triangulation of Δ. By this, we mean that Δ is partitioned into a set of simplices such that any two simplices intersect in either a common face, common vertex, or not at all. Let V be the vertex set of this triangulation. Note that any point p lies in some simplex in this triangulation, and, slightly abusing notation, let $V(p)$ be the set of corners for this simplex [5]. Informally, our rounding scheme rounds a point p to nearby points in V — p will be randomly mapped into $V(p)$ in the natural manner.

To formalize this, associate a test function $w_v(p)$ with each $v \in V$ as follows. Each distribution p can be uniquely written as a weighted average of its neighboring vertices, $V(p)$. For $v \in V(p)$, let us define the test functions $w_v(p)$ to be these linear weights, so they are uniquely defined by the linear equation:

$$p = \sum_{v \in V(p)} w_v(p) v\,.$$

For $v \notin V(p)$, we define $w_v(p) = 0$. A useful property is that

$$\sum_{v \in V(p)} w_v(p) = \sum_{v \in V} w_v(p) = 1$$

[4] The function g is Lipschitz continuous if g is continuous and if there exists a finite constant λ such that $|g(a) - g(b)| \leq \lambda|a - b|$.

[5] If this simplex is not unique, *ie* if p lies on a face, then choose any adjacent simplex

which holds since p is an average (under w_v) of the points in $V(p)$.

The functions w_v imply a natural randomized rounding function. Define the randomized rounding function $Round_V$ as follows: for some distribution p, $Round_V(p)$ chooses $v \in V(p)$ with probability $w_v(p)$. We make the following assumptions about a randomized rounding forecasting rule F_V with respect to F and triangulation V:

1. F is weakly calibrated.
2. If F makes the forecast f_t at time t, then F_V makes the random forecast $Round_V(f_t)$ at this time.
3. The (l_1) diameter of any simplex in the triangulation is less than ϵ, *ie* for any p and q in the same simplex, $|p - q| \leq \epsilon$.

An immediate corollary to the previous theorem is that F_V is ϵ-calibrated with respect to the indicator test functions.

Corollary 1. *For all X, the calibration error of F_V is asymptotically less than ϵ,* ie *the probability (taken with respect to the randomization used by $Round_V$) that*

$$|\mu_T(I_{p,\epsilon}, X, F_V)| \leq \epsilon$$

tends to 1 as T tends to infinity.

To see this, note that the instantaneous error at time t, $X_t - Round_V(f_t)$, has an expected value of $\sum_v w_v(f_t)(X_t - v)$ which is ϵ-close to $\sum_v w_v(f_t)(X_t - f_t)$. The sum of this latter quantity converges to 0 by the previous theorem. The (martingale) strong law of large numbers then suffices to prove this corollary.

This randomized scheme is "almost deterministic" in the sense that at each time t the forecast made by F_V is ϵ-close to a deterministic forecast. Interestingly, this shows that an adversarial nature cannot foil the forecaster, even if nature almost knows the forecast that will be used every round.

3 Publicly Calibrated Learning

First, some definitions are in order. Consider a game with n players. Each player i has a finite action space $\mathcal{A}_i$. The joint action space is then $\mathcal{A} = \Pi_{i=1}^n \mathcal{A}_i$. Associated with each player is a payoff function $u_i : \mathcal{A} \to [0, 1]$. The interpretation is that if the joint action $a \in \mathcal{A}$ is taken by all players then player i will receive payoff $u_i(a)$.

If p is a joint distribution over $\mathcal{A}_{-i} = \Pi_{j \neq i} \mathcal{A}_j$, then we define $BR_i(p)$ to be the set of all actions which are best responses for player i to p, *ie* it is the set of all $a \in \mathcal{A}_i$ which maximize the function $E_{a_{-i} \sim p}[u_i(a, a_{-i})]$. It is also useful to define ϵ-$BR_i(p)$ as the set of all actions which are ϵ-best responses to p, *ie* if $a \in \epsilon$-$BR_i(p)$ then the utility $E_{a_{-i} \sim p}[u_i(a, a_{-i})]$ is ϵ-close to the maximal utility $\max_{a' \in A} E_{a_{-i} \sim p}[u_i(a', a_{-i})]$.

Given some distribution f over $\mathcal{A}$, it is convenient to denote the marginal distribution of f over $\mathcal{A}_{-i}$ as f_{-i}. We say a distribution f is a Nash equilibrium (or, respectively, ϵ-Nash equilibrium) if the following two conditions hold:

1. f is a product distribution.
2. If action $a \in \mathcal{A}_i$ has positive probability under f then a is in $BR_i(f_{-i})$ (or, respectively, in ϵ-$BR_i(f_{-i})$).

We denote the set of all Nash equilibria (or ϵ-Nash equilibria) by NE (or NE_ϵ).

3.1 Using Public Forecasts

A standard setting for learning in games is for each player i to make some forecast p over $\mathcal{A}_{-i}$ at time t. The action taken by player i during this time would then be some action that is a best response to p.

Now consider the setting in which all players observe some forecast f_t over all n players, *ie* the forecast f_t is a full joint probability distribution over $\Omega = \mathcal{A}$. Each player is only interested in the prediction of other players, so player i can just use the marginal distribution $(f_t)_{-i}$ to form a prediction for the other players. In order to calibrate, some randomized rounding is in order.

More formally, we define the *public learning process* with respect to a forecasting rule F and vertex set V as follows: At each time t, F provides a prediction f_t and each player i:

1. makes a prediction $p = Round_V(f_t)$
2. chooses a best response to p_{-i}

We make the following assumptions.

1. F is weakly calibrated.
2. Ties for a best response are broken with a deterministic, stationary rule.
3. If p and q are in the same simplex (of the triangulation) then $|p - q| \leq \epsilon$.

It is straightforward to see that the forecasting rule of player i, which is $(Round_V(f_t))_{-i}$, is calibrated *regardless* of how the other players behave. By the previous corollary the randomized scheme $Round_V(f_t)$ will be ϵ-calibrated. Player i can then simply ignore the direction i of this forecast (by marginalizing) and hence has an ϵ-calibrated forecast over the reduced space $\mathcal{A}_{-i}$.

Thus, the rather minimal accuracy condition that players make calibrated predictions is satisfied, and, in this sense, it is rational for players to use the forecasts made by F. In fact, the setting of "publicly announced" forecasts is only one way to view the scheme. Alternatively, one could assume that each player has knowledge of the deterministic rule F and makes the computations of f_t themselves. Furthermore, F only needs the history of play as an input (and does not need any knowledge of the players' utility functions).

It is useful to make the following definitions. Let Convex(Q) be the set of all convex combinations of distributions in Q [6]. Define the distance between a distribution p and a set Q as:

$$d(p, Q) = \inf_{q \in Q} |p - q|$$

[6] If $q_1, q_2, \ldots q_m \in Q$ then $\alpha_1 q_1 + \alpha_2 q_2 \ldots + \alpha_m q_m \in \text{Convex}(Q)$, where α_i are positive and sum to one.

Using the result of Foster and Vohra [1997], it is immediate that the frequency of empirical play in the public learning process will (almost surely) converge into the set of 2ϵ-correlated equilibria, since the players are making ϵ-calibrated predictions, *ie*

$$d\left(\frac{1}{T}\sum_{t=1}^{T} X_t,\, CE_{2\epsilon}\right) \to 0$$

where $CE_{2\epsilon}$ is the set of 2ϵ-correlated equilibria. Our second main result shows we can further restrict the convergent set to convex combinations of Nash equilibria, a potentially much smaller set than the set of correlated equilibria.

Theorem 2. *(Nash Convergence) The joint frequency of empirical play in the public learning process converges into the set of convex combinations of 2ϵ-Nash equilibria,* ie *with probability* 1

$$d\left(\frac{1}{T}\sum_{t=1}^{T} X_t,\, Convex(NE_{2\epsilon})\right) \to 0$$

as T goes to infinity. Furthermore, the rule F rarely uses forecasts that are not close to a 2ϵ-Nash equilibrium — by this, we mean that with probability one

$$\frac{1}{T}\sum_{t=1}^{T} d(f_t,\, NE_{2\epsilon}) \to 0$$

as T goes to infinity.

Since our convergence is with respect to the *joint* empirical play, an immediate corollary is that the average payoff achieved by each player is at least the player's payoff under some 2ϵ-Nash equilibrium. Also, we have the following corollary showing convergence to NE.

Corollary 2. *If F is uniformly, weakly calibrated and if the triangulation V is made finer (*ie *if ϵ is decreased) sufficiently slowly, then the joint frequency of empirical play converges into the set of convex combinations of NE.*

As we stated in the Introduction, we argue that the above result deals with the coordination problem of "which Nash equilibrium to play?" in a sensible manner. Though the players cannot be pinned down to play any particular Nash equilibrium, they do jointly play some Nash equilibrium for long subsequences. Furthermore, it is public knowledge of which equilibrium is being played since the predictions f_t are frequently close to some Nash equilibrium (*not* general combinations of them).

Now of course if the Nash equilibrium is unique, then the empirical play converges to it. This does not contradict the (impossibility) result of Hart and Mas-Colell [2003] — crudely, our learning setting keeps track of richer statistics from the history of play (which is not permitted in their setting).

3.2 The Proof

On some round in which f is forecasted, every player acts according to a fixed randomized rule. Let $\pi(f)$ be this "play distribution" over joint actions $\mathcal{A}$ on any round with forecast f. More precisely, if f_t is the forecast at time t, then $\pi(f_t)$ is the expected value of X_t given f_t. Clearly, $\pi(f)$ is a product distribution since all players choose actions independently (since their randomization is private).

Lemma 1. *For all Lipschitz continuous test functions w, with probability 1, we have*

$$\frac{1}{\tau}\sum_{t=1}^{\tau} w(f_t)(f_t - \pi(f_t)) \to 0$$

as τ tends to infinity.

Proof. Consider the stochastic process $Y_\tau = \frac{1}{\tau}\sum_{t=1}^{\tau} w(f_t)(X_t - \pi(f_t))$. This is a martingale average (i.e. τY_τ is a martingale), since at every round, the expected value of X_t is $\pi(f_t)$. By the martingale strong law we have $Y_\tau \to 0$ as τ tends to infinity, with probability one. Also, by calibration, we have $\frac{1}{\tau}\sum_{t=1}^{\tau} w(f_t)(f_t - X_t) \to 0$ as τ tends to infinity. Combining these two leads to the result. □

We now show that fixed points of π are approximate Nash equilibria.

Lemma 2. *If $f = \pi(f)$, then f is a 2ϵ-Nash equilibrium.*

Proof. Assume that $a \in \mathcal{A}_i$ has positive probability under $\pi(f)$. By definition of the public learning process, action a must be a best response to some distribution p_{-i}, where $p \in V(f)$. Assumption 3 implies that $|p - f| \leq \epsilon$, so it follows that $|p_{-i} - f_{-i}| \leq \epsilon$. Since the utility of taking a under any distribution q_{-i} is $\sum_{a_{-i}\in\mathcal{A}_{-i}} q_{-i}[a_{-i}]u_i(a, a_{-i})$, the previous inequality and boundedness of u_i by 1 imply that a must be a 2ϵ-best response to f_{-i}. Furthermore, f is a product distribution, since $\pi(f)$ is one. The result follows. □

Taken together, these last two lemmas suggest that forecasts which are used often must be a 2ϵ-Nash equilibrium — the first lemma suggests that forecasts f which are used often must be equal to $\pi(f)$, and the second lemma states that if this occurs, then f is a 2ϵ Nash equilibrium. We now make this precise.

Define a forecast f to be *asymptotically unused* if there exists a continuous test function w such that $w(f) = 1$ and $\frac{1}{T}\sum_{t=1}^{T} w(f_t) \to 0$. In other words, a forecast is asymptotically unused if we can find some small neighborhood around it such that the limiting frequency of using a forecast in this neighborhood is 0.

Lemma 3. *If f is not a 2ϵ-Nash equilibrium, then it is asymptotically unused, with probability one.*

Proof. Consider a sequence of ever finer balls around f, and associate a continuous test function with each ball that is nonzero within the ball. Let r_1, r_2, r_3, ... be a sequence of decreasing radii such that $r_i \to 0$ as i tends to infinity. Define the open ball B_i as the set of all points p such that $|p - f| < r_i$. Associate a continuous test function w_i with the i-th ball such that: if $p \notin B_i$, $w_i(p) = 0$ and if $p \in B_i$, $w_i(p) > 0$, with $w_i(f) = 1$. Clearly, this construction is possible.

Define the radius r'_i as the maximal variation of π within the the $i-th$ ball, *ie* $r'_i = \sup_{p,q \in B_i} |\pi(p) - \pi(q)|$. Since $\pi(p)$ is continuous, then $r'_i \to 0$ as i tends to infinity.

Using the fact that $|f - \pi(f)|$ is a constant (for the following first equality),

$$
\begin{aligned}
&\left(\sum_{t=1}^{T} w_i(f_t)\right) |f - \pi(f)| \\
&= \left|\sum_{t=1}^{T} w_i(f_t)(f - \pi(f))\right| \\
&= \left|\sum_{t=1}^{T} w_i(f_t)\Big((f - f_t) - (\pi(f) - \pi(f_t)) + (f_t - \pi(f_t))\Big)\right| \\
&\leq \left|\sum_{t=1}^{T} w_i(f_t)(f - f_t)\right| + \left|\sum_{t=1}^{T} w_i(f_t)(\pi(f) - \pi(f_t))\right| + \left|\sum_{t=1}^{T} w_i(f_t)(f_t - \pi(f_t))\right| \\
&\leq (r_i + r'_i)\sum_{t=1}^{T} w_i(f_t) + \left|\sum_{t=1}^{T} w_i(f_t)(f_t - \pi(f_t))\right|
\end{aligned}
$$

where the last step uses the fact that $w_i(f_t)$ is zero if $|f_t - f| \geq r_i$ (*ie* if $f_t \notin B_i$) along with the definitions of r_i and r'_i.

Now to prove that f is asymptotically unused it suffices to show that there exists some i such that $\frac{1}{T}\sum_{t=1}^{T} w_i(f_t) \to 0$ as T tends to infinity. For a proof by contradiction, assume that such an i does not exist. Dividing the above equation by these sum weights, which are (asymptotically) nonzero by this assumption, we have

$$|f - \pi(f)| \leq r_i + r'_i + \frac{|\frac{1}{T}\sum_{t=1}^{T} w_i(f_t)(f_t - \pi(f_t))|}{\frac{1}{T}\sum_{t=1}^{T} w_i(f_t)}$$

Now by lemma 1, we know the numerator of the last term goes to 0. So, for all i, we have that $|f - \pi(f)| \leq r_i + r'_i$. By taking the limit as i tends to infinity, we have $|f - \pi(f)| = 0$. Thus f is a 2ϵ-Nash equilibrium by the previous lemma, which contradicts our assumption on f. □

We say a set of forecasts Q is asymptotically unused if there exists a continuous test function w such that $w(f) = 1$ for all $f \in Q$ and $\frac{1}{T}\sum_{t=1}^{T} w(f_t) \to 0$.

Lemma 4. *If Q is a compact set of forecasts such that every $f \in Q$ is not a 2ϵ-Nash equilibrium, then Q is asymptotically unused, with probability one.*

Proof. By the last lemma, we know that each $q \in Q$ is asymptotically unused. Let w_q be a test function which proves that q is asymptotically unused. Since w_q is continuous and $w_q(q) = 1$, there exists an open neighborhood around q in which w_q is strictly positive. Let $N(q)$ be this open neighborhood.

Clearly the set Q is covered by the (uncountable) union of all open neighborhoods $N(q)$, *ie* $Q \subset \cup_{q \in Q} N(q)$. Since Q is compact, every cover of Q by open sets has a finite subcover. In particular, there exists a finite sized set $C \subset Q$ such that $Q \subset \cup_{c \in C} N(c)$.

Let us define the test function $w = \frac{1}{|C|} \sum_{c \in C} w_c$. We use this function to prove that Q is asymptotically unused (we modify it later to have value 1 on Q). This function is continuous, since each w_c is continuous. Also, w is non-zero for all $q \in Q$. To see this, for every $q \in Q$ there exists some $c \in C$ such that $q \in N(c)$ since C is a cover, and this implies that $w_c(q) > 0$. Furthermore, for every $c \in C$, $\frac{1}{T}\sum_{t=1}^{T} w_c(f_t) \to 0$ with probability one and since $|C|$ is finite, we have that $\frac{1}{T}\sum_{t=1}^{T} w(f_t) \to 0$ with probability one.

Since Q is compact, w takes on its minimum value on Q. Let $\alpha = \min_{q \in Q} w(q)$, so $\alpha > 0$ since w is positive on Q. Hence, the function $w(q)/\alpha$ is at least 1 on Q. Now the function $w'(q) = \min\{w(q)/\alpha, 1\}$ is continuous, one on Q, and with probability one, $\frac{1}{T}\sum_{t=1}^{T} w'(f_t) \to 0$. Therefore, w' proves that Q is asymptotically unused. □

It is now straightforward to prove theorem 2. We start by proving that $\frac{1}{T}\sum_{t=1}^{T} d(f_t, NE_{2\epsilon}) \to 0$ with probability one. It suffices to prove that with probability one, for all $\delta > 0$ we have that

$$\frac{1}{T}\sum_{t=1}^{T} I\Big[d(NE_{2\epsilon}, f_t) \geq \delta\Big] \to 0$$

where I is the indicator function. Let Q_δ be the set of q such that $d(q, NE_{2\epsilon}) \geq \delta$. This set is compact, so each Q_δ is asymptotically unused. Let w_δ be the function which proves this. Since $w_\delta(f_t) \geq I\Big[d(NE_{2\epsilon}, f_t) \geq \delta\Big]$ (with equality on Q_δ), the above claim follows since $\frac{1}{T}\sum_{t=1}^{T} w_\delta(f_t) \to 0$.

Now let us prove that $d\left(\frac{1}{T}\sum_{t=1}^{T} X_t, \text{Convex}(NE_{2\epsilon})\right) \to 0$ with probability one. First, note that calibration implies $\frac{1}{T}\sum_{t=1}^{T} X_t \to \frac{1}{T}\sum_{t=1}^{T} f_t$ (just take w to be the constant test function to see this). Now the above statement directly implies that $\frac{1}{T}\sum_{t=1}^{T} f_t$ must converge into the set $Convex(NE_{2\epsilon})$.

4 A Deterministically Calibrated Algorithm

We now provide an algorithm that is uniformly, weakly calibrated for a constructive proof of theorem 1. For technical reasons, it is simpler to allow our algorithm to make forecasts which are not valid probability distributions — the forecasts lie in the expanded set $\widetilde{\Delta}$, defined as:

$$\widetilde{\Delta} = \{f : \sum_{k \in \Omega} f[k] = 1 \text{ and } f[k] \geq -\epsilon\}$$

so clearly $\Delta \subset \widetilde{\Delta}$, where Δ is the probability simplex in $\mathcal{R}^{|\Omega|}$. We later show that we can run this algorithm and simply project its forecasts back onto Δ (which does not alter our convergence results).

Similar to Subsection 2.2, consider a triangulation over this larger set $\widetilde{\Delta}$ with vertex set V, and let $V(p)$ be the corners of the simplex which contain p. It useful to make the following assumptions:

1. If p, q are in the same simplex in the triangulation, $|p - q| \leq \epsilon$.
2. Associated with each $v \in V$ we have a test function w_v which satisfies:
   (a) If $v \notin V(p)$, then $w_v(p) = 0$.
   (b) For all $p \in \widetilde{\Delta}$, $\sum_v w_v(p) = 1$ and $\sum_v w_v(p)v = p$.
3. For convenience, assume ϵ is small enough ($\epsilon \leq \frac{1}{4|\Omega|}$) suffices) such that for all $p, q \in \widetilde{\Delta}$, we have $|p - q| \leq 3$ (whereas for all $p, q \in \Delta$, $|p - q| \leq 2$).

In the first subsection, we present an algorithm, *Forecast the Fixed Point*, which (uniformly) drives the calibration error to 0 for those functions w_v. As advertised, the algorithm simply forecasts a fixed point of a particular function. It turns out that these fixed points can be computed efficiently (by tracking how the function changes at each timestep), but we do not discuss this here. The next subsection provides the analysis of this algorithm, which uses an "approachability" argument along with properties of the fixed point. Finally, we take $\epsilon \to 0$ which drives the calibration error to 0 (at a bounded rate) for any Lipschitz continuous test function, thus proving uniform, weak calibration.

4.1 The Algorithm: Forecast the Fixed Point

For notational convenience, we use $\mu_T(v)$ instead of $\mu_T(w_v, X, F)$, *ie*

$$\mu_T(v) = \frac{1}{T}\sum_{t=1}^{T} w_v(f_t)(X_t - f_t)$$

For $v \in V$, define a function $\rho_T(v)$ which moves v along the direction of calibration error $\mu_T(v)$, *ie*

$$\rho_T(v) = v + \mu_T(v)$$

For an arbitrary point $p \in \widetilde{\Delta}$, define $\rho_T(p)$ by interpolating on V. Since $p = \sum_{v \in V} w_v(p)v$, define $\rho_T(p)$ as:

$$\begin{aligned}\rho_T(p) &= \sum_{v \in V} w_v(p)\rho_T(v) \\ &= p + \sum_{v \in V} w_v(p)\mu_T(v)\end{aligned}$$

Clearly, this definition is consistent with the above when $p \in V$. In the following section, we show that ρ_T maps $\widetilde{\Delta}$ into $\widetilde{\Delta}$, which allows us to prove that ρ_T has a fixed point in $\widetilde{\Delta}$ (using Brouwer's fixed point theorem).

The algorithm, *Forecast the Fixed Point*, chooses a forecast $f \in \widetilde{\Delta}$ at time T which is any fixed point of the function ρ_{T-1}, *ie*:

1. At time $T = 1$, set $\mu_0(v) = 0$ for all $v \in V$.
2. At time T, compute a fixed point of ρ_{T-1}.
3. Forecast this fixed point.

4.2 The Analysis of This Algorithm

First, let us prove the algorithm exists.

Lemma 5. *(Existence) For all X and T, a fixed point of ρ_T exists in $\widetilde{\Delta}$. Furthermore, the forecast f_T at time T satisfies:*

$$\sum_{v \in V} w_v(f_T)\mu_{T-1}(v) = 0$$

Proof. We use Brouwer's fixed point theorem to prove existence, which involves proving that: 1) the mapping is into, *ie* $\rho_T : \widetilde{\Delta} \to \widetilde{\Delta}$ and 2) the mapping is continuous. First, let us show that $\rho_T(v) \in \widetilde{\Delta}$ for points $v \in V$. We know

$$\begin{aligned}\rho_T(v) &= v + \frac{1}{T}\sum_{t=1}^{T} w_v(f_t)(X_t - f_t) \\ &= \left(1 - \frac{1}{T}\sum_{t=1}^{T} w_v(f_t)\right) v + \frac{1}{T}\sum_{t=1}^{T} w_v(f_t)(X_t + v - f_t)\end{aligned}$$

It suffices to prove that $X_t + v - f_t$ is in $\widetilde{\Delta}$ (when $w_v(f_t) > 0$), since then the above would be in $\widetilde{\Delta}$ (by the convexity of $\widetilde{\Delta}$). Note that $w_v(f_t) = 0$ when $|v - f_t| > \epsilon$. Now if $|v - f_t| \le \epsilon$, then $v - f_t$ perturbs each component of X_t by at most ϵ, which implies that $X_t + v - f_t \in \widetilde{\Delta}$ since $X_t \in \Delta$. For general points $p \in \widetilde{\Delta}$, the mapping $\rho_T(p)$ must also be in Δ, since the mapping is an interpolation. The mapping is also continuous since the w_v's are continuous. Hence, a fixed point exists. The last equation follows by setting $\rho_{T-1}(f_T) = f_T$. □

Now let us bound the summed l_2 error, where $||x|| = \sqrt{x \cdot x}$.

Lemma 6. *(Error Bound) For any X, we have*

$$\sum_{v \in V} ||\mu_T(v)||^2 \le \frac{9}{T}$$

Proof. It is more convenient to work with the unnormalized quantity $r_T(v) = T\mu_T(v) = \sum_{t=1}^{T} w_v(f_t)(X_t - f_t)$. Note that

$$||r_T(v)||^2 = ||r_{T-1}(v)||^2 + w_v(f_T)^2||X_T - f_T||^2 + 2w_v(f_T)r_{T-1}(v) \cdot (X_T - f_T)$$

Summing the last term over V, we have

$$\begin{aligned}\sum_{v \in V} w_v(f_T)r_{T-1}(v) \cdot (X_T - f_T) &= T(X_T - f_T) \cdot \sum_{v \in V} w_v(f_T)\mu_{T-1}(v) \\ &= 0\end{aligned}$$

where we have used the fixed point condition of the previous lemma. Summing the middle term over V and using $||X_T - f_T|| \le |X_T - f_T| \le 3$, we have:

$$\begin{aligned}\sum_{v \in V} w_v(f_T)^2||X_T - f_T||^2 &\le 9\sum_{v \in V} w_v(f_T)^2 \\ &\le 9\sum_{v \in V} w_v(f_T) \\ &= 9\end{aligned}$$

Using these bounds along with some recursion, we have

$$\begin{aligned}\sum_{v\in V} ||r_T(v)||^2 &\leq \sum_{v\in V} ||r_{T-1}(v)||^2 + 9 \\ &\leq 9T\end{aligned}$$

The result follows by normalizing (*ie* by dividing the above by T^2). □

4.3 Completing the Proof for Uniform, Weak Calibration

Let g be an arbitrary Lipschitz function with Lipschitz parameter λ_g, *ie* $|g(a) - g(b)| \leq \lambda_g |a - b|$. We can use V to create an approximation of g as follows

$$\hat{g}(p) = \sum_{v\in V} g(v) w_v(p).$$

This is a good approximation in the sense that:

$$|\hat{g}(p) - g(p)| \leq \epsilon\lambda_g$$

which follows from the Lipschitz condition and the fact that $p = \sum_{v\in V} w_v(p)v$.

Throughout this section we let F be "Forecast the Fixed Point". Using the definition of $\mu_T(g, X, F)$ along with $|X_t - f_t| \leq 3$, we have

$$|\mu_T(g, X, F)| \leq \left| \frac{1}{T} \sum_{t=1}^{T} \hat{g}(f_t)(X_t - f_t) \right| + 3\epsilon\lambda_g = |\mu_T(\hat{g}, X, F)| + 3\epsilon\lambda_g$$

Continuing and using our shorthand notation of $\mu_T(v)$,

$$\begin{aligned}|\mu_T(\hat{g}, X, F)| &= \left| \frac{1}{T} \sum_{t=1}^{T} \sum_{v\in V} g(v) w_v(f_t)(X_t - f_t) \right| \\ &= \left| \sum_{v\in V} g(v) \mu_T(w_v, X, F) \right| \\ &\leq \sum_{v\in V} |\mu_T(v)| \\ &\leq \sqrt{|V| \sum_{v\in V} ||\mu_T(v)||^2}\end{aligned}$$

where the first inequality follows from the fact that $g(v) \leq 1$, and the last from the Cauchy-Schwarz inequality.

Using these inequalities along with lemma 6, we have

$$|\mu_T(g, X, F)| \leq \sqrt{\frac{9|V|}{T}} + 3\epsilon\lambda_g$$

Thus, for any fixed g we can pick ϵ small enough to kill off λ_g This unfortunately implies that $|V|$ is large (since the vertex set size grow with $1/\epsilon$). But we can make T large enough to kill off this $|V|$. To get convergence to precisely zero, we follow the usual approach of slowly tightening the parameters. This will be done in phases. Each phase will half the value of the target accuracy and will be long enough to cover the burn in part of the following phase (where error accrues).

Our proof is essentially complete, except for the fact that the algorithm F described so far could sometimes forecast outside the simplex (with probabilities greater than 1 or less than zero). To avoid this, we can project a forecast in $\widetilde{\Delta}$ onto the closest point in Δ. Let $P(\cdot)$ be such a projection operator. For any $f \in \widetilde{\Delta}$, we have $|P(f) - f| \leq |\Omega|\epsilon$. Thus, for any Lipschitz weighting function w we have

$$\begin{aligned} \mu_T(w, X, P \circ F) &= \sum_{v \in V} w(P(f_t))(X_t - P(f_t)) \\ &= \sum_{v \in V} w(P(f_t))(X_t - f_t) + \sum_{v \in V} w(P(f_t))(f_t - P(f_t)) \\ &\leq \mu_T(w \circ P, X, F) + |\Omega|\epsilon \end{aligned}$$

Hence the projected version also converges to 0 as $\epsilon \to 0$ (since $w \circ P$ is also Lipschitz continuous). Theorem 1 follows.

References

P. Billingsley. *Convergence of Probability Measures*. John Wiley and sons, 1968.

A. P. Dawid. The well-calibrated bayesian. *Journal of the Am. Stat. Assoc*, 77, 1982.

A. P. Dawid. The impossibility of inductive inference. *Journal of the Am. Stat. Assoc*, 80, 1985.

D. Foster and R. Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 1997.

Dean Foster and Rakesh V. Vohra. Asymptotic calibration. *Biometrika*, 85, 1998.

Dean P. Foster and H. Peyton Young. On the impossibility of predicting the behavior of rational agents. *Proceedings of the National Academy of Sciences*, 98, 2001.

Dean P. Foster and H. Peyton Young. Learning, hypothesis testing, and nash equilibrium. forthcoming.

Y. Freund and R. Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29, 1999.

D. Fudenberg and D. Levine. *The Theory of Learning in Games*. MIT Press, 1998.

D. Fudenberg and D. Levine. Conditional universal consistency. *Games and Economic Behavior*, 29:104 – 130, 1999.

S. Hart and A. Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68, 2000.

S. Hart and A. Mas-Colell. Uncoupled dynamics do not lead to nash equilibrium. *American Economic Review*, 93:1830–1836, 2003.

E. Kalai, E. Lehrer, and R. Smorodinsky. Calibrated forecasting and merging. *Games and Economic Behavior*, 29, 1999.

Ehud Kalai and Ehud Lehrer. Rational learning leads to nash equilibrium. *Econometrica*, 1993.

Paul Milgrom and John Roberts. Adaptive and sophisticated learning in normal form games. *Games and Economic Behavior*, 3:82 – 100, 1991.

D. Oakes. Self-calibrating priors do not exist. *Journal of the Am. Stat. Assoc*, 80, 1985.

J Robinson. An iterative method of solving a game. *Ann. Math.*, 1951.