
Calibration is a concept that tries to formalize a notion of quality for forecasters. For example, suppose a weatherman predicts each day whether the it will rain, or be sunny. Typically forecasters will predict such events in terms of probabilities, i.e., "There is a 30% chance of rain." Given only the outcome that day, it is impossible to judge the quality of such a forecast. However, if we consider *all* days on which a forecaster said the probability of rain was x%, it is reasonable to expect that the fraction of such days on which it rained is exactly x%. This is precisely the notion of calibration.

In this lecture we first define a notion of calibration that is appropriate for the study of games, and then prove that calibration is essentially a generalization of internal regret minimization. Formally, we will show that playing a best response to a calibrated forecast of the opponent is an internal regret minimizing strategy; and that using an internal regret minimizing algorithm, one can easily build a calibrated forecaster. Our presentation is based primarily on the corresponding paper of Foster and Vohra [2].

Throughout the lecture we consider a finite two-player game, where each player i has a finite pure action set A_i; let $A = \prod_i A_i$, and let $A_{-i} = \prod_{j \neq i} A_j$. We let a_i denote a pure action for player i, and let $s_i \in \Delta(A_i)$ denote a mixed action for player i. We will typically view s_i as a vector in $\mathbb{R}^{A_i}$, with $s_i(a_i)$ equal to the probability that player i places on a_i. We let $\Pi_i(\boldsymbol{a})$ denote the payoff to player i when the composite pure action vector is $\boldsymbol{a}$, and by an abuse of notation also let $\Pi_i(\boldsymbol{s})$ denote the expected payoff to player i when the composite mixed action vector is $\boldsymbol{s}$.

The game is played repeatedly by the players. We let $h^T = (\boldsymbol{a}^0, \ldots, \boldsymbol{a}^{T-1})$ denote the history up to time T. The *internal regret* of player i of action a_i against action a_i' after history h^T is:

$$IR_i(h^T; a_i, a_i') = \sum_{t=0}^{T-1} \mathcal{I}\{a_i^t = a_i\} \left(\Pi_i(a_i', \boldsymbol{a}_{-i}^t) - \Pi_i(a_i, \boldsymbol{a}_{-i}^t) \right).$$

We let $q^T \in \Delta(A_1 \times A_2)$ denote the joint empirical distribution of play play up to time $T - 1$:

$$q^T(a_1, a_2) = \frac{1}{T} \sum_{t=0}^{T-1} \mathcal{I}\{a_1^t = a_1, a_2^t = a_2\}.$$

1 Calibration

We assume that, at each time t, player 1 makes a *forecast* $f_{12}^t \in \Delta(A_2)$ of the mixed action that player 2 will play. We define $N(s_2, T)$ as the number of times that player 1 has forecast s_2 in the first T time steps:

$$N(s_2, T) = \sum_{t=0}^{T-1} \mathcal{I}\{f_{12}^t = s_2\}.$$

We define $\rho_1^T(a_2; s_2)$ as the fraction of time that player 2 played action a_2, among those time periods where player 1 forecast s_2, up to time $T-1$:

$$\rho_1^T(a_2; s_2) = \frac{\sum_{t=0}^{T-1} \mathcal{I}\{f_{12}^t = s_2, a_2^t = a_2\}}{\sum_{t=0}^{T-1} \mathcal{I}\{f_{12}^t = s_2\}}.$$

(Define $\rho_1^T(a_2; s_2) = 0$ if the denominator in the preceding expression is zero.) We say that the forecaster used by player 1 is *calibrated* if the following limit holds almost surely, regardless of the (possibly history-dependent) strategy used by player 2:

$$\lim_{T\to\infty} \sum_{s_2 \in \Delta(s_2)} \left|\rho_1^T(a_2; s_2) - s_2(a_2)\right| \left(\frac{N(s_2, T)}{T}\right) = 0, \quad \text{for all } a_2 \in A_2. \tag{1}$$

(Since player 1 has made only finitely many forecasts up to time T, the sum is well defined for all finite T.) Thus we look at the limiting fraction of time that player 2 plays a_2, when s_2 is forecast. Informally, on this subsequence of time periods, the fraction of time that player 2 plays a_2 must approach $s_2(a_2)$. The sum weighted by $N(s_2, T)/T$ ensures uniformity of calibration in the limit; i.e., the calibration error must approach zero uniformly over the forecasts chosen by player 1. (We note that many other, typically weaker, formulations of calibration are often used in the literature; we refer the reader to [1] for details.)

2 Calibration Implies Internal Regret Minimization

We start with the following simple theorem: best responses to a calibrated forecaster will minimize internal regret.

Theorem 1 *Suppose that player 1 uses a calibrated forecast of player 2's play, and at each time t plays a pure best response to this forecast; assume that ties are broken according to a stationary and deterministic tiebreaking rule. Then the resulting strategy for player 1 is internal regret minimizing.*

Proof. The proof idea is to note that calibration is a form of "internal regret minimization in forecast space." We will proceed by first grouping together all the forecasts that would lead to a given mixed action s_1 played by player 1, and then use this interpretation of calibration to establish internal regret minimization.

Formally, given $a_1 \in A_1$, let $F_1(a_1) \subset \Delta(A_2)$ be the set of mixed actions of player 2 for which a_1 is a best response (under the stationary, deterministic tiebreaking rule that has been chosen). Note that $\bigcup_{a_1 \in A_1} F_1(A_1) = \Delta(A_2)$. Further, note that $F_1(a_1) \subset BR_1^{-1}(a_1)$, where $BR_1 : \Delta(A_2) \to \Delta(A_1)$ is the best response map of player 1.

We start with the following calculation:

$$
\begin{aligned}
q^T(a_1,a_2) &= \frac{1}{T}\sum_{t=0}^{T-1}\mathcal{I}\{a_1^t = a_1, a_2^t = a_2\} \\
&= \frac{1}{T}\sum_{t=0}^{T-1}\mathcal{I}\{f_{12}^t \in F_1(a_1), a_2^t = a_2\} \\
&= \frac{1}{T}\sum_{t=0}^{T-1}\sum_{f\in F_1(a_1)}\mathcal{I}\{f_{12}^t = f, a_2^t = a_2\} \\
&= \frac{1}{T}\sum_{f\in F_1(a_1)}\rho_1^T(a_2;f)N(f,T) \\
&= \frac{1}{T}\sum_{f\in F_1(a_1)} f(a_2)N(f,T) + \\
&\quad \sum_{f\in F_1(a_1)}\left(\rho_1^T(a_2;f) - f(a_2)\right)\left(\frac{N(f,T)}{T}\right).
\end{aligned}
$$

The first equality follows by definition of the joint empirical distribution. The second equality uses the fact that player 1 plays a_1 if and only if the forecast lies in $F_1(a_1)$. The fourth equality uses the definitions of ρ_1^T and N.

In the last equality, notice that the second summation converges to zero almost surely as $T \to \infty$, by the assumption of calibrated forecasting. Thus we have:

$$
\begin{aligned}
\frac{1}{T}IR_1(h^T;a_1,a_1') &= \frac{1}{T}\sum_{t=0}^{T-1}\mathcal{I}\{a_1^t = a_1\}\left(\Pi_1(a_1',a_2^t) - \Pi_1(a_1,a_2^t)\right) \\
&= \sum_{a_2\in A_2} q^T(a_1,a_2)(\Pi(a_1',a_2) - \Pi(a_1,a_2)) \\
&leq \frac{1}{T}\sum_{a_2\in A_2}\sum_{f\in F_1(a_1)} f(a_2)N(f,T)(\Pi(a_1',a_2) - \Pi(a_1,a_2)) + \varepsilon_T \\
&= \sum_{f\in F_1(a_1)}\left(\frac{N(f,T)}{T}\right)\sum_{a_2\in A_2} f(a_2)(\Pi(a_1',a_2) - \Pi(a_1,a_2)) + \varepsilon_T \\
&= \sum_{f\in F_1(a_1)}\left(\frac{N(f,T)}{T}\right)(\Pi(a_1',f) - \Pi(a_1,f)) + \varepsilon_T \\
&\leq \varepsilon_T,
\end{aligned}
$$

where ε_T is an error term that approaches zero as $T \to \infty$. The first equality is the definition of internal regret. The second equality follows by rewriting the first expression. The first inequality follows by our expression for q^T in terms of calibration error. The remaining equalities follow by rearranging terms. Finally, the last inequality follows since for every $f \in F_1(a_1)$, a_1 is a best response for player 1; thus $\Pi(a_1', f) - \Pi(a_1, f) \leq 0$.

From the preceding we conclude that for all a_1, a_1', there holds almost surely:

$$\limsup_{T\to\infty} \frac{1}{T} IR_1(h^T; a_1, a_1') \leq 0,$$

as required. □

Remarks:

1. The original paper of Foster and Vohra actually establishes that if all players use the suggested algorithm (via a calibrated forecaster), then play converges to the set of correlated equilibria. Of course, this is a trivial consequence of the preceding result, since play converges to the set of correlated equilibria if all players use internal regret minimizing algorithms.

2. Note that best response to the calibrated forecaster is a form of "fictitious play." On other hand, note that the marginal empirical distribution of the opponent need not be a calibrated forecast, so standard fictitious play is (obviously) not calibrated. (We know this already, since standard fictitious play need not even minimize external regret.) More generally, it is not hard to show that no deterministic forecaster can be calibrated.

3 Internal Regret Minimization implies ε-Calibration

We now show that internal regret minimizing algorithms can be used to build calibrated forecasters, establishing a form of equivalence between the two concepts (taken together with the last section). For simplicity, we focus attention on the case where player 2 has only two actions available: $A_2 = \{0, 1\}$. (This is also called the problem of *binary sequence prediction*.) In addition, rather than asking for exact calibration, we only establish the weaker notion of ε-calibration; this is not an enormous limitation, as it is possible to show (through an application of the doubling trick) that the family of ε-calibrated forecasters we build can be used to build a single calibrated forecaster.

Since we are only predicting binary sequences, we interpret a forecast as the probability that the next play of player 2 will be 1; thus a forecast is a real number in $[0, 1]$. Using notation analogous to the preceding section, define $\rho_1^T(p)$ and $N(p, T)$ as follows:

$$N(p, T) = \sum_{t=0}^{T-1} \mathcal{I}\{f_{12}^t = p\}; \quad \rho_1^T(p) = \frac{\sum_{t=0}^{T-1} a_2^t \mathcal{I}\{f_{12}^t = p\}}{\sum_{t=0}^{T-1} \mathcal{I}\{f_{12}^t = p\}}.$$

Given $\varepsilon > 0$, we seek a forecaster that is ε-calibrated, i.e., that satisfies the following almost surely, regardless of the (possibly history-dependent) strategy of player 2:

$$\limsup_{T\to\infty} \sum_{p \in [0,1]} \left|\rho_1^T(p) - p\right| \left(\frac{N(p, T)}{T}\right) \leq \varepsilon.$$

Again, the sum is well defined for every finite T.

Our main idea is to *discretize* the forecast space. Fix a positive integer k, and suppose that forecasts are chosen only from the set $F = \{0, 1/k, 2/k, \ldots, (k-1)/k, 1\}$. We consider a "forecasting game", i.e., a game where the loss to player 1 when a forecast $f \in F$ is made and player 2 plays $a \in \{0, 1\}$ is:

$$\ell(f, a) = (a - f)^2.$$

(Equivalently, the payoff to player 1 is $-\ell(f, a)$.) Player 1's goal in this game is to minimize his loss. We will show that if player 1 uses an internal regret minimizing algorithm in the forecasting game, and k is large enough, then the resulting forecaster is ε-calibrated.

This approach yields the following theorem.

Theorem 2 *Given a strategy of player 2, let f_{12}^t denote the forecast (in F) chosen at time t by an internal regret minimizing algorithm in the forecasting game defined above. Then regardless of the strategy of player 2, the resulting sequence of forecasts $\{f_{12}^t\}$ is calibrated.*

Proof. We first note that it suffices to show:

$$\limsup_{T\to\infty} \sum_{p\in[0,1]} (\rho_1^T(p) - p)^2 \left(\frac{N(p,T)}{T}\right) \le \varepsilon^2.$$

This follows by using Jensen's inequality and the definition of calibration.

Let w_1^t denote the mixed action over F played by player 1 at time t in the forecasting game, according to the internal regret minimizing algorithm. (Note that this is *a mixed action over forecasts*!) For $p \in F$, define:

$$\overline{\rho}_1^T(p) = \frac{\sum_{t=0}^{T-1} a_2^t w_1^t(p)}{\sum_{t=0}^{T-1} w_1^t(p)}.$$

We start by showing that, in an appropriate sense, actual play can be replaced by expectations. Using a standard argument (via the Azuma-Hoeffding inequality and the Borel-Cantelli Lemma), it follows that for all $p \in F$ (almost surely):

$$\lim_{T\to\infty} \frac{1}{T} \left| \sum_{t=0}^{T-1} a_2^t w_1^t(p) - \sum_{t=0}^{T-1} a_2^t \mathcal{I}\{f_{12}^t = p\} \right| = 0,$$

and

$$\lim_{T\to\infty} \frac{1}{T} \left| \sum_{t=0}^{T-1} w_1^t(p) - \sum_{t=0}^{T-1} \mathcal{I}\{f_{12}^t = p\} \right| = 0.$$

From these two limits, it follows also that:

$$\lim_{T\to\infty} \left| \sum_{p\in F} (\rho_1^T(p) - p)^2 \left(\frac{N(p,T)}{T}\right) - (\overline{\rho}_1^T(p) - p)^2 \left(\frac{\sum_{t=0}^{T-1} w_1^t(p)}{T}\right) \right| = 0.$$

Thus it suffices to show that:

$$\limsup_{T\to\infty} \sum_{p\in F} (\overline{\rho}_1^T(p) - p)^2 \left(\frac{\sum_{t=0}^{T-1} w_1^t(p)}{T} \right) \leq \varepsilon.$$

Note that:

$$\mathbb{E}\left[\mathcal{I}\left\{ f_{12}^t = \frac{i}{k} \right\} \left(\ell\left(\frac{i}{k}, a_2^t\right) - \ell\left(\frac{j}{k}, a_2^t\right) \right) \right] = w_1^t\left(\frac{i}{k}\right) \left(\left(a_2^t - \frac{i}{k}\right)^2 - \left(a_2^t - \frac{j}{k}\right)^2 \right).$$

(The expectation is only with respect to player 1's randomization.)

Noting that $(a_2^t - i/k)^2 - (a_2^t - j/k)^2 = (2a_2^t - i/k - j/k)(j/k - i/k)$, we have:

$$\begin{aligned}
\mathbb{E}\left[IR_1\left(h^T; \frac{i}{k}, \frac{j}{k}\right) \right] &= \sum_{t=0}^{T-1} \mathbb{E}\left[\mathcal{I}\left\{ f_{12}^t = \frac{i}{k} \right\} \left(\ell\left(\frac{i}{k}, a_2^t\right) - \ell\left(\frac{j}{k}, a_2^t\right) \right) \right] \\
&= \sum_{t=0}^{T-1} w_1^t\left(\frac{i}{k}\right) \left(\frac{j-i}{k}\right) \left(2a_2^t - \frac{i+j}{k}\right) \\
&= \left(\sum_{t=0}^{T-1} w_1^t\left(\frac{i}{k}\right) \right) \left(\frac{j-i}{k}\right) \left(2\overline{\rho}_1^T\left(\frac{i}{k}\right) - \frac{i+j}{k}\right) \\
&= \left(\sum_{t=0}^{T-1} w_1^t\left(\frac{i}{k}\right) \right) \left(\left(\overline{\rho}_1^T\left(\frac{i}{k}\right) - \frac{i}{k}\right)^2 - \left(\overline{\rho}_1^T\left(\frac{i}{k}\right) - \frac{j}{k}\right)^2 \right).
\end{aligned}$$

Minimizing the right hand side over j, note that we can always choose j/k to be within at most $1/k$ of $\overline{\rho}_1^T(i/k)$, since F is a $1/k$-discretization of $[0,1]$. Thus there exists at least one choice of j for which we have:

$$\left(\frac{\sum_{t=0}^{T-1} w_1^t(i/k)}{T} \right) \left(\overline{\rho}_1^T\left(\frac{i}{k}\right) - \frac{i}{k}\right)^2 \leq \frac{1}{T}\mathbb{E}\left[IR_1\left(h^T; \frac{i}{k}, \frac{j}{k}\right) \right] + \frac{1}{k^2}.$$

If we sum over i on the left hand side and take the sup over i on the right hand side, we obtain:

$$\sum_{p\in F} (\overline{\rho}_1^T(p) - p)^2 \left(\frac{\sum_{t=0}^{T-1} w_1^t(p)}{T} \right) \leq (k+1) \left(\sup_{p,p'\in F} \frac{1}{T}\mathbb{E}[IR_1(h^T; p, p')] \right) + (k+1)/k^2,$$

since i ranges from 0 to k. Choosing k sufficiently large (in particular, so that $(k+1)/k^2 < \varepsilon^2$), and taking $T \to \infty$ yields the desired result. $\square$

Remarks:

1. The choice of loss function matters in the proof. For example, trying to prove the result using standard L^1 loss will fail (i.e., directly defining the loss as the absolute forecasting error). You are encouraged to check this for yourself; see also [1] for further details.

2. It is worth emphasizing that in the preceding result, the regret bound scales linearly with the number of forecasts in the discretization. In particular, if player 2 has general finite action set A_2, the number of points in an ε-discretization of $\Delta(A_2)$ scales *exponentially* in the size of A_2. Thus calibration is ultimately creating a virtual forecasting game in which the action space of player 1 is siginificantly expanded, and then applying internal regret minimization in that space. On the other hand, the application of calibration to establish internal regret minimization in the preceding section amounts to a *compression* of the forecast space, by grouping together forecasts that lead to the same (pure) best response by player 1.

References

[1] N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, Cambridge, United Kingdom, 2004.

[2] D. Foster and R. Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21:40–55, 1997.