# Prediction under limited feedback

**Gábor Lugosi**

Pompeu Fabra University, Barcelona

based on joint work with

**Nicolò Cesa-Bianchi and Gilles Stoltz**

## Motivating example: on-line pricing

A vendor sells $n$ pieces of a product to $n$ customers.

Customers come one by one.

To customer number $t$, the vendor offers the product at a price $I_t \in [0, 1]$.

Each customer has a maximum price $y_t$ he/she is willing to pay but does not tell it to the vendor.

If $y_t \geq I_t$, the product is bought and the vendor suffers a "loss" $y_t - I_t$.

If $y_t < I_t$, the product is not bought and the vendor's loss is $c \in [0, 1]$.

# On-line pricing

The vendor's loss function is

$$\ell(I_t, y_t) = (y_t - I_t)\mathbb{I}_{I_t \le y_t} + c\mathbb{I}_{I_t > y_t}$$

The values $y_t$ are arbitrary and may even depend on the vendor's past actions.

If the vendor knew the "distribution" of $y_1, \ldots, y_n$ in advance, he could choose the value $p$ minimizing the total loss

$$\frac{1}{n}\sum_{t=1}^{n} \ell(p, y_t) \ .$$

**Result:** The vendor has a (randomized) strategy such that, for all $\delta \in [0,1]$, with probability at least $1-\delta$,

$$\begin{aligned} & \frac{1}{n}\sum_{t=1}^{n} \ell(I_t, y_t) - \min_{p\in[0,1]} \frac{1}{n}\sum_{t=1}^{n} \ell(p, y_t) \\ \le \quad & Cn^{-1/5}\sqrt{\ln(n/\delta)} \ . \end{aligned}$$

# Randomized prediction

A game between forecaster and environment.

At each round $t$, the forecaster chooses an action $I_t \in \{1, \ldots, N\}$;

the environment chooses an action $y_t \in \mathcal{Y}$;

the forecaster suffers loss $\ell(I_t, y_t) \in [0, 1]$.

The goal is to minimize the *cumulative excess loss*

$$\frac{1}{n}\left(\sum_{t=1}^{n} \ell(I_t, y_t) - \min_{i \leq N} \sum_{t=1}^{n} \ell(i, y_t)\right) .$$

The forecaster may randomize. At time $t$ chooses a probability distribution $\mathbf{p}_t = (p_{1,t}, \ldots, p_{N,t})$

and plays action $i$ with probability $p_{i,t}$.

Actions are often called "experts".

## Randomized prediction

This and related models have been studied in

- game theory: playing repeated games;
- information theory: gambling and data compression;
- statistics: sequential decisions;
- statistical learning theory: on-line learning;

The simplest model assumes that after each round, the losses $\ell(i, y_t)$ $(i = 1, \ldots, N)$ are revealed *(full information)*.

In this model **Hannan** (1957) showed that the forecaster has a strategy such that

$$\frac{1}{n}\left(\sum_{t=1}^{n} \ell(I_t, y_t) - \min_{i \le N} \sum_{t=1}^{n} \ell(i, y_t)\right) \to 0$$

almost surely for all strategies of the environment.

# Hannan consistency: basic ideas

Obviously, the forecaster must randomize.

$$\ell(\mathbf{p}_t, y_t) = \sum_{i=1}^{N} p_{i,t}\ell(i, y_t) = \mathbf{E}_t \ell(I_t, y_t)$$

denotes the "expected" loss of the forecaster.

By martingale convergence,

$$\frac{1}{n}\left(\sum_{t=1}^{n} \ell(I_t, y_t) - \sum_{t=1}^{n} \ell(\mathbf{p}_t, y_t)\right) = O_P(n^{-1/2})$$

so it suffices to study

$$\frac{1}{n}\left(\sum_{t=1}^{n} \ell(\mathbf{p}_t, y_t) - \min_{i \le N} \sum_{t=1}^{n} \ell(i, y_t)\right)$$

## Weighted average prediction

Idea: assign a higher probability to better-performing actions.

A popular choice is

$$p_{i,t} = \frac{\exp\left(-\eta \sum_{s=1}^{t} \ell(i,y_s)\right)}{\sum_{k=1}^{N} \exp\left(-\eta \sum_{s=1}^{t} \ell(k,y_s)\right)} \quad i = 1,\ldots,N\,.$$

where $\eta > 0$. Then

$$\begin{aligned}\frac{1}{n}\left(\sum_{t=1}^{n} \ell(\mathbf{p}_t,y_t) - \min_{i\le N}\sum_{t=1}^{n} \ell(i,y_t)\right) \quad &\le \quad \frac{\ln N}{n\eta} + \frac{\eta}{8} \\ &= \quad \sqrt{\frac{\ln N}{2n}}\end{aligned}$$

with $\eta = \sqrt{8\ln N/n}$.

# Proof

Let $L_{i,t} = \sum_{s=1}^{t} \ell(i, y_s)$ and

$$W_t = \sum_{i=1}^{N} w_{i,t} = \sum_{i=1}^{N} e^{-\eta L_{i,t}}$$

for $t \geq 1$, and $W_0 = N$. First observe that

$$\begin{aligned}
\ln \frac{W_n}{W_0} &= \ln \left( \sum_{i=1}^{N} e^{-\eta L_{i,n}} \right) - \ln N \\
&\geq \ln \left( \max_{i=1,\ldots,N} e^{-\eta L_{i,n}} \right) - \ln N \\
&= -\eta \min_{i=1,\ldots,N} L_{i,n} - \ln N \ .
\end{aligned}$$

## Proof

On the other hand, for each $t = 1, \ldots, n$

$$\begin{aligned}
\ln \frac{W_t}{W_{t-1}} &= \ln \frac{\sum_{i=1}^N w_{i,t-1} e^{-\eta \ell(i,y_t)}}{\sum_{j=1}^N w_{j,t-1}} \\
&\le -\eta \frac{\sum_{i=1}^N w_{i,t-1} \ell(i,y_t)}{\sum_{j=1}^N w_{j,t-1}} + \frac{\eta^2}{8} \\
&= -\eta \ell(\mathbf{p}_t, y_t) + \frac{\eta^2}{8}
\end{aligned}$$

by Hoeffding's inequality.

Summing over $t = 1, \ldots, n$,

$$\ln \frac{W_n}{W_0} \le -\eta \sum_{t=1}^n \ell(\mathbf{p}_t, y_t) + \frac{\eta^2}{8} n \ .$$

Combining these, we get

$$\sum_{t=1}^n \ell(\mathbf{p}_t, y_t) \le \min_{i=1,\ldots,N} L_{i,n} + \frac{\ln N}{\eta} + \frac{\eta}{8} n$$

# Lower bound

The upper bound is optimal in the sense that for all predictors,

$$\sup_{n,N,y_1^n} \frac{\sum_{t=1}^n \ell(I_t, y_t) - \min_{i \le N} \sum_{t=1}^n \ell(i, y_t)}{\sqrt{(n/2)\ln N}} \ge 1 \ .$$

(Cesa-Bianchi, Freund, Haussler, Helmbold, Schapire, and Warmuth'97).

Idea: choose $\ell(i, y_t)$ to be i.i.d. symmetric Bernoulli.

Then the best predictor is random guessing.

Use the central limit theorem.

# Label efficient prediction

Model proposed by Helmbold and Panizza (1997).

In this variant the forecaster does not see the outcome $y_t$ unless he asks for it, but can do it only $m \ll n$ times.

The game is the following:

For each round $t = 1, \ldots, n$,

(1) the environment chooses the outcome $y_t \in \mathcal{Y}$ without revealing it;

(2) the forecaster chooses $\mathbf{p}_t$ and draws an action $I_t \in \{1, \ldots, N\}$ according to this distribution;

(3) the forecaster incurs loss $\ell(I_t, y_t)$ and each action $i$ incurs loss $\ell(i, y_t)$, none of these values is revealed to the forecaster;

(4) the forecaster decides whether he asks for the value of $y_t$ if the total number of revealed outcomes up to time $t - 1$ is less than $m$.

# A label efficient forecaster

The idea is to ask for labels randomly (with probability $\approx m/n$) and use the weighted average forecaster with the estimated losses.

Let $Z_t$ be i.i.d. Bernoulli $\epsilon = \frac{m - \sqrt{2m\ln(4/\delta)}}{n}$ and $\eta = \sqrt{\frac{2\varepsilon \ln N}{n}}$.

The forecaster asks for $y_t$ iff $Z_t = 1$.

Let

$$\widetilde{\ell}(i, y_t) \stackrel{\text{def}}{=} \begin{cases} \ell(i, y_t)/\varepsilon & \text{if } Z_t = 1, \\ 0 & \text{otherwise.} \end{cases}$$

An unbiased estimate!

For each round $t = 1, 2, \ldots, n$ draw an action from $\{1, \ldots, N\}$ according to the distribution

$$p_{i,t} = \frac{\exp\left(-\eta \sum_{s=1}^{t} \widetilde{\ell}(i, y_s)\right)}{\sum_{k=1}^{N} \exp\left(-\eta \sum_{s=1}^{t} \widetilde{\ell}(k, y_s)\right)} \quad i = 1, \ldots, N\,.$$

# Bound for label efficient prediction

With probability at least $1 - \delta$,

$$\frac{1}{n}\left(\sum_{t=1}^{n} \ell(I_t, y_t) - \min_{i \le N} \sum_{t=1}^{n} \ell(i, y_t)\right)$$
$$\le 9\sqrt{\frac{\ln N + \ln(4/\delta)}{m}} \ .$$

Sketch of proof:

First bound

$$\sum_{t=1}^{n} \widetilde{\ell}(\mathbf{p}_t, y_t) - \min_{i \le N} \sum_{t=1}^{n} \widetilde{\ell}(i, y_t)$$

by standard methods. Then use Bernstein-type martingale inequalities to handle

$$\sum_{t=1}^{n} \ell(I_t, y_t) - \sum_{t=1}^{n} \widetilde{\ell}(\mathbf{p}_t, y_t)$$

and

$$\min_{i \le N} \sum_{t=1}^{n} \widetilde{\ell}(i, y_t) - \min_{i \le N} \sum_{t=1}^{n} \ell(i, y_t)$$

# Hannan consistency

There exists a randomized label efficient forecaster that achieves Hannan consistency while issuing, for all $n > 1$, at most $O((\ln \ln n)^2 \ln n)$ queries in the first $n$ prediction steps.

Consistency is achieved with only a logarithmic number of labels!

We don't know if this rate is optimal.

Proof: divide time into consecutive blocks of length $1, 2, 4, 8, 16, \ldots$.

In the $r$-th block use the forecaster with parameters $n = 2^{r-1}$, $m = (\ln r)(\ln \ln r)$ and $\delta = 1/r^3$.

# Lower bound

There exists a loss function $\ell$ such that for all sets of $N$ actions any forecaster asking for at most $m$ labels has

$$\sup_{y_1,\ldots,y_n\in\{0,1\}} \mathbf{E}\frac{1}{n}\left(\sum_{t=1}^{n}\ell(I_t,y_t)-\min_{i\le N}\sum_{t=1}^{n}\ell(i,y_t)\right)$$
$$\ge c\sqrt{\frac{\ln N}{m}}\ .$$

Idea (for $N=2$): choose the outcomes randomly (i.i.d.) such that they are either Bernoulli $1/2-\epsilon$ or Bernoulli $1/2+\epsilon$.

Interpretation: $m$ acts like a sample size.

A similar phenomenon occurs in the multi-armed bandit problem.

## Improved bound for small losses

Let $L_n^* = \min_{i\le N} \sum_{t=1}^n \ell(i, y_t)$

If the forecaster is used with parameters $\epsilon = \frac{m-\sqrt{2m\ln(4/\delta)}}{n}$ and $\eta = \min\left\{\sqrt{\frac{2\epsilon \ln N}{L_n^*}}, \epsilon\right\}$, then

$$\frac{1}{n}\left(\sum_{t=1}^n \ell(I_t, y_t) - \min_{i\le N}\sum_{t=1}^n \ell(i, y_t)\right)$$
$$\le \quad C\left(\sqrt{\frac{L_n^*}{n}\frac{1}{m}\log\frac{N}{\delta}} + \frac{1}{m}\log\frac{N}{\delta}\right)$$

# Partial monitoring

In the most general setup the information received by the forecaster after making a prediction $I_t$ is a *feedback* $h(I_t, y_t)$.

We assume that $\mathcal{Y} = \{1, \ldots, M\}$.

The matrix of losses is $\mathbf{L} = [\ell(i,j)]_{N \times M}$ (known by the forecaster).

At time $t$, the forecater chooses action $I_t \in \{1, \ldots, N\}$ and the outcome is $y_t \in \mathcal{Y}$.

The forecaster's loss is $\ell(I_t, y_t)$.

The forecaster only observes the feedback $h(I_t, y_t)$ where $\mathbf{H} = [h(i,j)]_{N \times M}$ is the feedback matrix (with values from a finite set).

# Prediction with partial monitoring

For each round $t = 1, \ldots, n$,

(1) the environment chooses the next outcome $y_t \in \mathcal{Y}$ without revealing it;

(2) the forecaster chooses a probability distribution $\mathbf{p}_t$ and draws an action $I_t \in \{1, \ldots, N\}$ according to $\mathbf{p}_t$;

(3) the forecaster incurs loss $\ell(I_t, y_t)$ and each action $i$ incurs loss $\ell(i, y_t)$. None of these values is revealed to the forecaster;

(4) the feedback $h(I_t, y_t)$ is revealed to the forecaster.

# Examples

**Dynamic pricing.** Here $M = N$, and $\mathbf{L} = [\ell(i,j)]_{N \times N}$ where

$$\ell(i,j) = \frac{(j-i)\mathbb{I}_{i \le j} + c\mathbb{I}_{i>j}}{N} \ .$$

and $h(i,j) = \mathbb{I}_{i>j}$ or

$$h(i,j) = a\mathbb{I}_{i \le j} + b\mathbb{I}_{i>j} \ , \quad i,j = 1,\ldots,N \ .$$

**Multi-armed bandit problem.** The only information the forecaster receives is his own loss: $\mathbf{H} = \mathbf{L}$.

## Examples

**Apple tasting.** $N = M = 2$.

$$\mathbf{L} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \mathbf{H} = \begin{bmatrix} a & a \\ b & c \end{bmatrix} .$$

The predictor only receives feedback when he chooses the second action. (Helmbold, Littlestone, and Long, 2000)

**Label efficient prediction.** $N = 3$, $M = 2$.

$$\mathbf{L} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\mathbf{H} = \begin{bmatrix} a & b \\ c & c \\ c & c \end{bmatrix} .$$

# A general predictor

A forecaster first proposed by Piccolboni and Schindelhauer.

Crucial assumption: $\mathbf{H}$ can be encoded such that there exists an $N \times N$ matrix $\mathbf{K} = [k(i,j)]_{N\times N}$ such that

$$\mathbf{L} = \mathbf{K} \cdot \mathbf{H} \; .$$

Thus,

$$\ell(i,j) = \sum_{l=1}^{N} k(i,l)h(l,j) \; .$$

Then we may estimate the losses by

$$\widetilde{\ell}(i,y_t) = \frac{k(i,I_t)h(I_t,y_t)}{p_{I_t,t}} \; .$$

# A general predictor

Observe

$$\begin{aligned}\mathbf{E}_t \widetilde{\ell}(i, y_t) &= \sum_{k=1}^{N} p_{k,t} \frac{k(i,k)h(k,y_t)}{p_{k,t}} \\ &= \sum_{k=1}^{N} k(i,k)h(k,y_t) = \ell(i,y_t) \ ,\end{aligned}$$

$\widetilde{\ell}(i, y_t)$ is an unbiased estimate of $\ell(i, y_t)$.

Let

$$p_{i,t} = (1-\gamma) \frac{e^{-\eta \widetilde{L}_{i,t-1}}}{\sum_{k=1}^{N} e^{-\eta \widetilde{L}_{k,t-1}}} + \frac{\gamma}{N}$$

where $\widetilde{L}_{i,t} = \sum_{s=1}^{t} \widetilde{\ell}(i, y_t)$.

# Performance bound

For all $\delta \in [0,1]$, with probability at least $1-\delta$,

$$
\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n}\ell(I_t,y_t) - \min_{i=1,\ldots,N}\frac{1}{n}\sum_{t=1}^{n}\ell(i,y_t) \\
\leq \quad Cn^{-1/3}N^{2/3}\sqrt{\ln(N/\delta)}\ .
\end{aligned}
$$

where $C$ depends on $\mathbf{K}$.

Thus, Hannan consistency is achieved with rate $O(n^{-1/3})$ whenever $\mathbf{L} = \mathbf{K}\cdot\mathbf{H}$.

This solves the dynamic pricing problem.

Extends to random feedbacks.

Whenever Hannan consistency is achievable, a version of this predictor works and attains rate $O(n^{-1/3})$.

## Revealing actions

Sometimes $\mathbf{L}$ can not be expressed as $\mathbf{K} \cdot \mathbf{H}$, yet Hannan consistency is achievable. Example:

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{and} \qquad \mathbf{H} = \begin{bmatrix} a & b & c \\ d & d & d \\ e & e & e \end{bmatrix} .$$

Here playing the first action reveals the outcome $y_t$. and the label-efficient prediction algorithm may be used.

An action $i \in \{1, \ldots, N\}$ is *revealing* if all entries in the $i$-th row of the feedback matrix are different. In such a case we have

$$\frac{1}{n}\left(\sum_{t=1}^{n} \ell(I_t, y_t) - \min_{i=1,\ldots,N} L_{1,n}\right) \le 9n^{-1/3}\left(\ln\frac{5N}{\delta}\right)^{1/3} .$$

## Distinguishing actions

Assume that $\mathbf{H}$ is such that for each outcome $j = 1, \ldots, M$ there exists an action $i \in \{1, \ldots, N\}$ such that for all outcomes $j' \neq j$, $h(i,j) \neq h(i,j')$.

Then there is a Hannan consistent forecaster with a normalized cumulative regret of the order of $n^{-1/3}$.

# Optimality

The example of label efficient prediction shows that the rate $O(n^{-1/3})$ is not improvable, in general.

**Bandit problems.** In this case $\mathbf{H} = \mathbf{L}$ so $\mathbf{K}$ is the identity matrix.

The forecaster becomes

$$p_{i,t} = (1-\gamma)\frac{e^{-\eta \widetilde{L}_{i,t-1}}}{\sum_{k=1}^{N} e^{-\eta \widetilde{L}_{k,t-1}}} + \frac{\gamma}{N}$$

suggested by Auer, Cesa-Bianchi, Freund, and Schapire.

They show that a carefully modified version achieves a faster $O(n^{-1/2})$ rate (as in the full information case).

## Questions

Give an attractive description of when Hannan consistency is achievable. (Piccolboni and Schindelhauer give a complicated procedure.)

Whenever Hannan consistency is achievable, one can have a rate of $n^{-1/3}$.

In some nontrivial cases $n^{-1/2}$ is achievable (e.g., bandit problem).

Another example: (a version of dynamic pricing) $M = N$,

$$\ell(i,j) = \frac{(j-i)\mathbb{I}_{i \le j} + (i-j)\mathbb{I}_{i>j}}{N} = \frac{|i-j|}{N}$$

and $h(i,j) = \mathbb{I}_{i>j}$.

Characterize the class of problems with fast rates. Is there any other rate?

Beyond Hannan consistency? Rustichini, Mannor and Shimkin.

# Further reading

N. Cesa-Bianchi and G. Lugosi.
*Prediction, Learning, and Games.*
Cambridge University Press ($\leq 2005 + \log(1/\delta)$)