Calibration, falsifiability and Macau

Dean Foster

Amazon.com, NYC
Setting: On-line decision making
(aka adversarial data or robust time series)
Goal: Use economic forecasts for decision making
Setting: On-line decision making 
(aka adversarial data or robust time series)

Goal: Use economic forecasts for decision making

Problem: Accuracy doesn’t guarantee good decisions
(We’ll take “accuracy” = “low regret.” Regret compares actual decisions to “20/20 hindsight.” 100s of papers say how to get low regret.)
Setting: On-line decision making
(aka adversarial data or robust time series)

Goal: Use economic forecasts for decision making

Problem: Accuracy doesn’t guarantee good decisions
(We’ll take “accuracy” = “low regret.” Regret compares actual decisions to “20/20 hindsight.” 100s of papers say how to get low regret.)

Solution: Falsifiable is better of error
- you falsify a forecast by betting against it
- The amount it loses is its macau.
My message in one slide

- Setting: On-line decision making
  \((aka \text{ adversarial data or robust time series})\)
- Goal: Use economic forecasts for decision making
- Problem: Accuracy doesn’t guarantee good decisions
  \((We’ll \text{ take } "accuracy" = "low \text{ regret}". \text{ Regret compares actual decisions to } "20/20 \text{ hindsight}". 100s of papers say how to get low regret.\)
- Solution: Falsifiable is better of error
  - you falsify a forecast by betting against it
  - The amount it loses is its \textit{macau}.

\textbf{Take Aways}

\textit{crazy-Calibration} + \textit{low-regret} \implies \textit{low-macau} \implies \textit{good decisions}
Prove the Earth is round!

Fun question: What personal evidence do you have that the earth is round?
Fun question: What personal evidence do you have that the earth is round?

Can you prove it is round? NO!

But, you can make claims that could easily be shown wrong.

Called falsifiability
We will falsify someone’s claim by winning bets placed against them.

Claim: $\hat{Y} \approx EY$

- Prove it wrong by winning lots of money:

\[
\text{expected winnings} = E \left( B (Y - \hat{Y}) \right)
\]

- $(Y - \hat{Y})$ is a “fair” bet
- $B$ is amount bet
We will falsify someone’s claim by winning bets placed against them

Claim: \( \hat{Y} \approx EY \)

- Prove it wrong by winning lots of money:
  
  \[
  \text{expected winnings} = E \left( B (Y - \hat{Y}) \right)
  \]

- \( Y - \hat{Y} \) is a “fair” bet
- \( B \) is amount bet

How to avoid being proven wrong by:

\[
E \left( B (Y - \hat{Y}) \right)
\]

(Start with bet \( B \))
Operationalizing falsifiability

- We will falsify someone’s claim by winning bets placed against them.
- Claim: $\hat{Y} \approx EY$
  - Prove it wrong by winning lots of money:
    \[
    \text{expected winnings} = E \left( B (Y - \hat{Y}) \right)
    \]
  - $(Y - \hat{Y})$ is a “fair” bet
  - $B$ is amount bet
- How to avoid being proven wrong by:
  \[
  \text{Macau} \equiv \max_{|B| \leq 1} E \left( B (Y - \hat{Y}) \right)
  \]
  (worry about worst bet)
Operationalizing falsifiability

- We will falsify someone’s claim by winning bets placed against them
- Claim: $\hat{Y} \approx EY$
  - Prove it wrong by winning lots of money:
    
    \[
    \text{expected winnings} = E \left( B (Y - \hat{Y}) \right)
    \]
  - $(Y - \hat{Y})$ is a “fair” bet
  - $B$ is amount bet
- How to avoid being proven wrong by:
  
  \[
  \min_{\hat{Y}} \max_{|B| \leq 1} E \left( B (Y - \hat{Y}) \right)
  \]
  
  \text{(mini-max)}
On to calibration

Average Predicted

Average Y

0.1 0.15 0.2 0.25

0.3 0.4 0.5

Average Predicted

0.1 0.15 0.2 0.25
<table>
<thead>
<tr>
<th>$Y$</th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_1$</td>
<td>$X_{11}$</td>
<td>$X_{12}$</td>
<td>$X_{13}$</td>
<td>$X_{14}$</td>
</tr>
<tr>
<td>$Y_2$</td>
<td>$X_{21}$</td>
<td>$X_{22}$</td>
<td>$X_{23}$</td>
<td>$X_{24}$</td>
</tr>
<tr>
<td>$Y_3$</td>
<td>$X_{31}$</td>
<td>$X_{32}$</td>
<td>$X_{33}$</td>
<td>$X_{34}$</td>
</tr>
<tr>
<td>$Y_4$</td>
<td>$X_{41}$</td>
<td>$X_{42}$</td>
<td>$X_{43}$</td>
<td>$X_{44}$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$Y_t$</td>
<td>$X_{t1}$</td>
<td>$X_{t2}$</td>
<td>$X_{t3}$</td>
<td>$X_{t4}$</td>
</tr>
</tbody>
</table>

Starting with our data that we observed up to time $t$
We can fit $\hat{\beta}_t$ on everything up to time $t$.
Crazy calibration variable

<table>
<thead>
<tr>
<th></th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_1$</td>
<td>$X_{11}$</td>
<td>$X_{12}$</td>
<td>$X_{13}$</td>
<td>$X_{14}$</td>
</tr>
<tr>
<td>$Y_2$</td>
<td>$X_{21}$</td>
<td>$X_{22}$</td>
<td>$X_{23}$</td>
<td>$X_{24}$</td>
</tr>
<tr>
<td>$Y_3$</td>
<td>$X_{31}$</td>
<td>$X_{32}$</td>
<td>$X_{33}$</td>
<td>$X_{34}$</td>
</tr>
<tr>
<td>$Y_4$</td>
<td>$X_{41}$</td>
<td>$X_{42}$</td>
<td>$X_{43}$</td>
<td>$X_{44}$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$Y_t$</td>
<td>$X_{t1}$</td>
<td>$X_{t2}$</td>
<td>$X_{t3}$</td>
<td>$X_{t4}$</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
Y_t + 1 &= \beta_t' X_{t+1} + \epsilon_t + \beta_t' X_t \\
\hat{Y}_{t+1} &= \hat{\beta}_t' X_{t+1}
\end{align*}

From a new $X_{t+1}$ we can compute $\hat{Y}_{t+1}$
Looking at only the first part of the data, we can generate:

$$\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4, \ldots, \hat{\beta}_{t-1}$$
Crazy calibration variable

\[
\begin{array}{cccccc}
Y & X_1 & X_2 & X_3 & X_4 & \beta \\
Y_1 & X_{11} & X_{12} & X_{13} & X_{14} & 0 \\
Y_2 & X_{21} & X_{22} & X_{23} & X_{24} & \beta_1 \\
Y_3 & X_{31} & X_{32} & X_{33} & X_{34} & \beta_2 \\
Y_4 & X_{41} & X_{42} & X_{43} & X_{44} & \beta_3 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
Y_t & X_{t1} & X_{t2} & X_{t3} & X_{t4} & \beta_{t-1} \\
\end{array}
\]

\[
\hat{Y} = \hat{\beta}' X_t
\]

Each of these leads to a next round

\[
\hat{Y}_1, \hat{Y}_2, \hat{Y}_3, \hat{Y}_4, \ldots, \hat{Y}_t
\]
Crazy calibration variable

<table>
<thead>
<tr>
<th>$Y$</th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$\hat{\beta}$</th>
<th>$\hat{Y}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_1$</td>
<td>$X_{11}$</td>
<td>$X_{12}$</td>
<td>$X_{13}$</td>
<td>$X_{14}$</td>
<td>0</td>
<td>$\hat{Y}_1 = 0$</td>
</tr>
<tr>
<td>$Y_2$</td>
<td>$X_{21}$</td>
<td>$X_{22}$</td>
<td>$X_{23}$</td>
<td>$X_{24}$</td>
<td>$\hat{\beta}_1$</td>
<td>$\hat{Y}_2 = \hat{\beta}_1' X_2$</td>
</tr>
<tr>
<td>$Y_3$</td>
<td>$X_{31}$</td>
<td>$X_{32}$</td>
<td>$X_{33}$</td>
<td>$X_{34}$</td>
<td>$\hat{\beta}_2$</td>
<td>$\hat{Y}_3 = \hat{\beta}_2' X_3$</td>
</tr>
<tr>
<td>$Y_4$</td>
<td>$X_{41}$</td>
<td>$X_{42}$</td>
<td>$X_{43}$</td>
<td>$X_{44}$</td>
<td>$\hat{\beta}_3$</td>
<td>$\hat{Y}_4 = \hat{\beta}_3' X_4$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$Y_t$</td>
<td>$X_{t1}$</td>
<td>$X_{t2}$</td>
<td>$X_{t3}$</td>
<td>$X_{t4}$</td>
<td>$\hat{\beta}_{t-1}$</td>
<td>$\hat{Y}<em>t = \hat{\beta}</em>{t-1}' X_t$</td>
</tr>
</tbody>
</table>

Theorem (Foster 1991, Forster 1999)

**Such an on-line least squares forecast generates low regret:**

$$\sum_{t=1}^{T} (Y_t - \hat{Y}_t)^2 - \min_{\beta} \sum_{t=1}^{T} (Y_t - \beta' X_t)^2 \leq O(\log(T))$$
Crazy calibration variable

\begin{tabular}{|c|cccc|}
\hline
$Y$ & $X_1$ & $X_2$ & $X_3$ & $X_4$ \\
\hline
$Y_1$ & $X_{11}$ & $X_{12}$ & $X_{13}$ & $X_{14}$ \\
$Y_2$ & $X_{21}$ & $X_{22}$ & $X_{23}$ & $X_{24}$ \\
$Y_3$ & $X_{31}$ & $X_{32}$ & $X_{33}$ & $X_{34}$ \\
$Y_4$ & $X_{41}$ & $X_{42}$ & $X_{43}$ & $X_{44}$ \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
$Y_t$ & $X_{t1}$ & $X_{t2}$ & $X_{t3}$ & $X_{t4}$ \\
\hline
\end{tabular}

\begin{align*}
\hat{\beta} & \quad \hat{Y} \\
0 & \quad \hat{Y}_1 = 0 \\
\hat{\beta}_1 & \quad \hat{Y}_2 = \hat{\beta}_1' X_2 \\
\hat{\beta}_2 & \quad \hat{Y}_3 = \hat{\beta}_2' X_3 \\
\hat{\beta}_3 & \quad \hat{Y}_4 = \hat{\beta}_3' X_4 \\
\vdots & \quad \vdots \\
\hat{\beta}_{t-1} & \quad \hat{Y}_t = \hat{\beta}_{t-1}' X_t \\
\end{align*}

Works no matter what the $X$'s are.
Crazy calibration variable

<table>
<thead>
<tr>
<th>$Y$</th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$\hat{\beta}$</th>
<th>$\hat{Y}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_1$</td>
<td>$X_{11}$</td>
<td>$X_{12}$</td>
<td>$\hat{Y}_1$</td>
<td>$X_{14}$</td>
<td>0</td>
<td>$\hat{Y}_1 = 0$</td>
</tr>
<tr>
<td>$Y_2$</td>
<td>$X_{21}$</td>
<td>$X_{22}$</td>
<td>$\hat{Y}_2$</td>
<td>$X_{24}$</td>
<td>$\hat{\beta}_1$</td>
<td>$\hat{Y}_2 = \hat{\beta}'_1 X_2$</td>
</tr>
<tr>
<td>$Y_3$</td>
<td>$X_{31}$</td>
<td>$X_{32}$</td>
<td>$\hat{Y}_3$</td>
<td>$X_{34}$</td>
<td>$\hat{\beta}_2$</td>
<td>$\hat{Y}_3 = \hat{\beta}'_2 X_3$</td>
</tr>
<tr>
<td>$Y_4$</td>
<td>$X_{41}$</td>
<td>$X_{42}$</td>
<td>$\hat{Y}_4$</td>
<td>$X_{44}$</td>
<td>$\hat{\beta}_3$</td>
<td>$\hat{Y}_4 = \hat{\beta}'_3 X_4$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$Y_t$</td>
<td>$X_{t1}$</td>
<td>$X_{t2}$</td>
<td>$\hat{Y}_t$</td>
<td>$X_{t4}$</td>
<td>$\hat{\beta}_{t-1}$</td>
<td>$\hat{Y}<em>t = \hat{\beta}'</em>{t-1} X_t$</td>
</tr>
</tbody>
</table>

Even if one of the $X$’s were $\hat{Y}$!
Crazy calibration variable

Theorem (\(\implies\) Foster and Kakade 2008, Foster and Hart 2018)

Adding the crazy calibration variable generates low macau:

\[
(\forall i) \sum_{t=1}^{T} X_{t,i} (Y_t - \hat{Y}_t) = O(\sqrt{T \log(T)})
\]
### Macau as the “normal equation”

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Least squares</th>
<th>Normal equations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(Y</td>
<td>X)$</td>
<td>$\min_\beta \sum (Y_i - \beta \cdot X_i)^2$</td>
</tr>
</tbody>
</table>

**The normal equation is the same as:**

$$\max_\alpha \sum_i \alpha' X_i(Y_i - \beta' X_i) = 0$$

**Which is solved by the $\beta$ minimizer:**

$$\min_\beta \max_\alpha \sum_i \alpha' X_i(Y_i - \beta' X_i) = 0$$
Macau as the “normal equation”

| $E(Y|X)$ | Least squares | Normal equations |
|----------|---------------|-----------------|
| Statistics | $\min_\beta \sum (Y_i - \beta \cdot X_i)^2$ | $\min_\beta \max_\alpha \sum \alpha \cdot X_i \ (Y_i - \beta \cdot X_i)$ |
### Macau as the “normal equation”

<table>
<thead>
<tr>
<th></th>
<th>Least squares</th>
<th>Normal equations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statistics</strong></td>
<td>[ \min_{\beta} \sum (Y_i - \beta \cdot X_i)^2 ]</td>
<td>[ \min_{\beta} \max_{\alpha} \sum \alpha \cdot X_i (Y_i - \beta \cdot X_i) ]</td>
</tr>
<tr>
<td><strong>Probability</strong></td>
<td>[ \min_{f} E((Y - f(X))^2) ] aka ( E(Y</td>
<td>X) )</td>
</tr>
</tbody>
</table>

The normal equation is the same as:

\[ \max_g E(g(X)(Y - f(X))) = 0 \]

Which is solved by the \( f(\cdot) \) minimizer:

\[ \min_{f} \max_g E(g(X)(Y - f(X))) = 0 \]
### Macau as the “normal equation”

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Least squares</th>
<th>Normal equations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(Y</td>
<td>X)$</td>
<td>$\min_{\beta} \sum (Y_i - \beta \cdot X_i)^2$</td>
</tr>
<tr>
<td>Probability</td>
<td>$\min_{f} E((Y - f(X))^2)$</td>
<td>$\min_{f} \max_{g} E\left(g(X) \ (Y - f(X))\right)$</td>
</tr>
</tbody>
</table>

aka $E(Y|X)$
Macau as the “normal equation”

<table>
<thead>
<tr>
<th></th>
<th>Least squares</th>
<th>Normal equations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistics</td>
<td>$\min_{\beta} \sum_i (Y_i - \beta \cdot X_i)^2$</td>
<td>$\min_{\beta} \max_{\alpha} \sum_i \alpha \cdot X_i \ (Y_i - \beta \cdot X_i)$</td>
</tr>
<tr>
<td>Probability</td>
<td>$\min_f E((Y - f(X))^2)$</td>
<td>$\min_f \max_g E\left(g(X) \ (Y - f(X))\right)$</td>
</tr>
<tr>
<td>online</td>
<td>low regret</td>
<td>low macau</td>
</tr>
</tbody>
</table>

Regret $\equiv \sum_{t=1}^{T} (Y_t - \hat{Y}_t)^2 - \min_{\beta} \sum_{t=1}^{T} (Y_t - \beta \cdot X_t)^2$
## Macau as the “normal equation”

### Least squares

**Statistics**

\[ \min_{\beta} \sum (Y_i - \beta \cdot X_i)^2 \]

**Normal equations**

\[ \min_{\beta} \max_{\alpha} \sum \alpha \cdot X_i \ (Y_i - \beta \cdot X_i) \]

### Probability

**Statistics**

\[ \min_{f} E((Y - f(X))^2) \equiv \min_{f} \max_{\alpha} E(g(X) \ (Y - f(X))) \]

aka \( E(Y|X) \)

**Probability**

\[ \min_{f} \max_{g} E \left( g(X) (Y - f(X)) \right) \]

### Online

**Low regret**

**Low macau**

\[ \text{Macau} \equiv \max_{\alpha: |\alpha| \leq 1} \sum_{t=1}^{T} \alpha \cdot X_t \ (Y_t - \hat{Y}_t) \]
Macau as the “normal equation”

<table>
<thead>
<tr>
<th></th>
<th>Least squares</th>
<th>Normal equations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statistics</strong></td>
<td>( \min_\beta \sum (Y_i - \beta \cdot X_i)^2 )</td>
<td>( \min_\beta \max_\alpha \sum \alpha \cdot X_i \ (Y_i - \beta \cdot X_i) )</td>
</tr>
<tr>
<td><strong>Probability</strong></td>
<td>( \min_f E( (Y - f(X))^2 ) ) aka ( E(Y</td>
<td>X) )</td>
</tr>
<tr>
<td><strong>online</strong></td>
<td>low regret</td>
<td>low macau</td>
</tr>
</tbody>
</table>

- **Statistics:** Least squares \( \iff \) normal equations
- **Probability:** Least squares \( \iff \) normal equations
Macau as the “normal equation”

<table>
<thead>
<tr>
<th></th>
<th>Least squares</th>
<th>Normal equations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statistics</strong></td>
<td>$\min_{\beta} \sum (Y_i - \beta \cdot X_i)^2$</td>
<td>$\min_{\beta} \max_{\alpha} \sum \alpha \cdot X_i \ (Y_i - \beta \cdot X_i)$</td>
</tr>
<tr>
<td><strong>Probability</strong></td>
<td>$\min_{f} E((Y - f(X))^2)$</td>
<td>$\min_{f} \max_{g} E\left(g(X) \ (Y - f(X))\right)$</td>
</tr>
<tr>
<td>online</td>
<td>low regret</td>
<td>low macau</td>
</tr>
</tbody>
</table>

### Take Aways

on-line low regret $\iff$ on-line low macau
No regret $\iff$ not falsified

<table>
<thead>
<tr>
<th>$t$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>$\cdots$</th>
<th>T</th>
<th>T+1</th>
<th>T+2</th>
<th>T+3</th>
<th>$\cdots$</th>
<th>3T</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_t$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
<td>1</td>
</tr>
<tr>
<td>$X_t$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
<td>1</td>
</tr>
<tr>
<td>$\hat{Y}_t$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
<td>0</td>
<td>$\frac{1}{7}$</td>
<td>$\frac{2}{T-1}$</td>
<td>$\frac{3}{T+2}$</td>
<td>$\cdots$</td>
</tr>
</tbody>
</table>

Not falsified $\iff$ no regret

<table>
<thead>
<tr>
<th>$t$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>$\cdots$</th>
<th>T</th>
<th>T+1</th>
<th>$\cdots$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_t$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>$\cdots$</td>
<td>0</td>
<td>1</td>
<td>$\cdots$</td>
</tr>
<tr>
<td>$X_t$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
</tr>
<tr>
<td>$\hat{Y}_t$</td>
<td>$\frac{2}{7}$</td>
<td>$\frac{4}{7}$</td>
<td>$\frac{2}{7}$</td>
<td>$\frac{4}{7}$</td>
<td>$\cdots$</td>
<td>$\frac{6}{7}$</td>
<td>$\frac{4}{7}$</td>
<td>$\cdots$</td>
</tr>
</tbody>
</table>

How about a bet?

- Macau is zero
- Regret is $T/9$
- So: low macau $\iff$ low regret
No regret $\not\iff$ not falsified

<table>
<thead>
<tr>
<th>$t$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>$\cdots$</th>
<th>T</th>
<th>T+1</th>
<th>T+2</th>
<th>T+3</th>
<th>$\cdots$</th>
<th>3T</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_t$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
<td>0</td>
</tr>
<tr>
<td>$x_t$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
<td>1</td>
</tr>
<tr>
<td>$\hat{\gamma}_t$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
<td>$\frac{1}{T}$</td>
<td>$\frac{2}{T+1}$</td>
<td>$\frac{3}{T+2}$</td>
<td>$\cdots$</td>
<td>$\frac{2}{3}$</td>
</tr>
</tbody>
</table>

Not falsified $\not\iff$ no regret

<table>
<thead>
<tr>
<th>$t$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>$\cdots$</th>
<th>T</th>
<th>T+1</th>
<th>$\cdots$</th>
<th>T+1</th>
<th>$\cdots$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_t$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>$\cdots$</td>
<td>0</td>
<td>1</td>
<td>$\cdots$</td>
<td>1</td>
<td>$\cdots$</td>
</tr>
<tr>
<td>$x_t$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
<td>1</td>
<td>$\cdots$</td>
</tr>
<tr>
<td>$\hat{\gamma}_t$</td>
<td>.6</td>
<td>.4</td>
<td>.6</td>
<td>.4</td>
<td>$\cdots$</td>
<td>.6</td>
<td>.4</td>
<td>$\cdots$</td>
<td>.6</td>
<td>$\cdots$</td>
</tr>
</tbody>
</table>

- Macau is zero
- Regret is $T/9$
- So: low macau $\not\iff$ low regret

(Skipping these proofs)
Action $A$ makes $X$ dollars, action $B$ makes $Y$ dollars

- We want forecasts that are close to $X$ and $Y$
- We want to be close on average
- We will use least squares to estimate $X$ and $Y$

But, we want to take actions

Will good estimates of $X$ and $Y$ lead to good decisions about $A$ vs $B$?
Some notation:

\[ a = \text{action taken } \in \mathbb{R}^k \text{(eg inventory levels)} \]
\[ X_t = \text{Context at time } t \]
\[ a_t^* = \text{best action at time } t \]
\[ r_t(a) = \text{Reward at time } t \text{ playing } a \]
\[ V_t^* = \max_a E(r_t(a)|X_t) = E(r_t(a^*)|X_t) \]
\[ q_t(a) \leq E(r_t(a)|X_t) \leq \bar{q}_t(a) \]
Some notation:

\[ a = \text{action taken } \in \mathbb{R}^k \text{ (eg inventory levels)} \]
\[ X_t = \text{Context at time } t \]
\[ a_t^* = \text{best action at time } t \]
\[ r_t(a) = \text{Reward at time } t \text{ playing } a \]
\[ V_t^* = \max_a E(r_t(a)|X_t) = E(r_t(a^*)|X_t) \]
\[ q_t(a) \leq E(r_t(a)|X_t) \leq \overline{q}_t(a) \]

What are good falsifiable claims about \( a^* \)?
Contextual Bandits

Some notation:

\[
\begin{align*}
    a &= \text{action taken} \in \mathbb{R}^k (\text{eg inventory levels}) \\
    X_t &= \text{Context at time } t \\
    a_t^* &= \text{best action at time } t \\
    r_t(a) &= \text{Reward at time } t \text{ playing } a \\
    V_t^* &= \max_a E(r_t(a)|X_t) = E(r_t(a^*)|X_t) \\
    q_t(a) &\leq E(r_t(a)|X_t) \leq \bar{q}_t(a)
\end{align*}
\]

Too precise:
“Here are two bounding functions \( q \) and \( \bar{q} \):

\[
\begin{align*}
    q_t(a) &= \bar{q}_t(a)
\end{align*}
\]
Some notation:

\[ a = \text{action taken } \in \mathbb{R}^k \text{(eg inventory levels)} \]
\[ X_t = \text{Context at time } t \]
\[ a_t^* = \text{best action at time } t \]
\[ r_t(a) = \text{Reward at time } t \text{ playing } a \]
\[ V_t^* = \max_a E(r_t(a)|X_t) = E(r_t(a^*)|X_t) \]
\[ q_t(a) \leq E(r_t(a)|X_t) \leq \bar{q}_t(a) \]

Too loose:

- “Here is \( a_t^* \).”
Some notation:

\[ a = \text{action taken } \in \mathbb{R}^k \text{(eg inventory levels)} \]
\[ X_t = \text{Context at time } t \]
\[ a_t^* = \text{best action at time } t \]
\[ r_t(a) = \text{Reward at time } t \text{ playing } a \]
\[ V_t^* = \max_a E(r_t(a)|X_t) = E(r_t(a^*)|X_t) \]
\[ q_t(a) \leq E(r_t(a)|X_t) \leq \bar{q}_t(a) \]

Just right:

“Here is a target \( V^* \) and approximating quadratics around \( a^* \):

\[ \bar{q}_t(a) = V_t^* - q||a - a_t^*||^2 \]
\[ \bar{q}_t(a) - \underline{q}_t(a) = \Delta||a - a_t^*||^2 \]"
Why is low macau useful?

Supposed each $c_t(\cdot)$ is convex
Goal: play $a$ to minimize $C(a)$
Eg: We could use SGD on $\nabla c_t()$
called “on-line convex optimization” with regret:

$$\text{regret} \equiv \sum_{t=1}^{T} (c_t(\hat{a}_t) - c_t(a^*))$$
Why is low macau useful?

\[
C(a) = \sum_{t=1}^{T} c_t(a) \quad a^* \equiv \arg\min_a C(a)
\]

The regret is bounded by the gradient:

\[
\text{regret} \leq \sum_{t=1}^{T} (\hat{a}_t - a^*) \cdot \nabla c_t(\hat{a}_t)
\]
Why is low macau useful?

\[
C(a) = \sum_{t=1}^{T} c_t(a) \quad a^* \equiv \arg \min_a C(a)
\]

The regret is bounded by the gradient:

\[
\text{regret} = \sum_{t=1}^{T} (c_t(\hat{a}_t) - c_t(a^*))
\leq \sum_{t=1}^{T} (\hat{a}_t - a^*) \cdot \nabla c_t(\hat{a}_t)
= \sum_{t=1}^{T} (\hat{a}_t - a^*) \cdot \left( \nabla c_t(\hat{a}_t) - \hat{\nabla} c_t(\hat{a}_t) \right) + (\hat{a}_t - a^*) \cdot \hat{\nabla} c_t(\hat{a}_t)
\]
Why is low macau useful?

\[
C(a) = \sum_{t=1}^{T} c_t(a) \quad a^* \equiv \arg \min_a C(a)
\]

The regret is bounded by the gradient:

\[
\text{regret} = \sum_{t=1}^{T} (c_t(\hat{a}_t) - c_t(a^*))
\leq \sum_{t=1}^{T} (\hat{a}_t - a^*) \cdot \nabla c_t(\hat{a}_t)
\leq \sum_{t=1}^{T} (\hat{a}_t - a^*) \cdot \left( \nabla c_t(\hat{a}_t) - \nabla c_t(\hat{a}_t) \right) + (\hat{a}_t - a^*) \cdot \nabla c_t(\hat{a}_t)
\]

\[
\text{regret} \leq \text{macau!}
\]

(zero @ \hat{a}_t)
Why is low macau useful?

\[
C(a) = \sum_{t=1}^{T} c_t(a) \quad a^* \equiv \arg \min_a C(a)
\]

The regret is bounded by the gradient:

\[
\text{regret} = \sum_{t=1}^{T} (c_t(\hat{a}_t) - c_t(a^*)) \\
\leq \sum_{t=1}^{T} (\hat{a}_t - a^*) \cdot \nabla c_t(\hat{a}_t) \\
= \sum_{t=1}^{T} (\hat{a}_t - a^*) \cdot \left( \nabla c_t(\hat{a}_t) - \hat{\nabla} c_t(\hat{a}_t) \right) + (\hat{a}_t - a^*) \cdot \hat{\nabla} c_t(\hat{a}_t)
\]

\[
\text{regret} \leq \text{macau}
\]
without crazy-calibration variable
Using the crazy-calibration variable
Theorem (F. and Kakade 2008, new)

Let $R$ be the quadratic regret of a forecast $\hat{Y}_t$ against a linear regression on $X_t$. Let $M$ be the Macau of $\hat{Y}_t$ using linear functions of $X_t$ to create falsifying bets. Then if $\hat{Y}_t = [X_t]_0$, we have $R = o(T)$ iff $M = o(T)$. 
Calibration Theorem

**Theorem ( \( \implies \) F. and Kakade 2008, \( \iff \) new)**

Let \( R \) be the quadratic regret of a forecast \( \hat{Y}_t \) against a linear regression on \( X_t \). Let \( M \) be the Macau of \( \hat{Y}_t \) using linear functions of \( X_t \) to create falsifying bets. Then if \( \hat{Y}_t = [X_t]_0 \), we have \( R = o(T) \) iff \( M = o(T) \).

Note: Typically, \( R = O(\log(T)) \) iff \( M = \tilde{O}(\sqrt{T}) \) for the actual algorithms I know.

*(Sasha Rakhlin and Dylan Foster have a proof for IID.)*
Theorem (⇒ F. and Kakade 2008, ⇐ new)

Let $R$ be the quadratic regret of a forecast $\hat{Y}_t$ against a linear regression on $X_t$. Let $M$ be the Macau of $\hat{Y}_t$ using linear functions of $X_t$ to create falsifying bets. Then if $\hat{Y}_t = [X_t]_0$, we have $R = o(T)$ iff $M = o(T)$.

Proof sketch: Consider the forecasts $(1 - w)\hat{Y}_t + w\alpha \cdot X_t$ for the *any* $\alpha$. Let $Q(w)$ be the total quadratic error of this family of forecast. The following are equivalent:

- $Q(0) \leq Q(w)$ (No regret condition)
- $Q'(0)$ is zero. (No macau condition)
Recipe for good decisions

- List bets that you would make to show \( \hat{a}_t \) is not optimal
- Convert these to regression variables
- Add the crazy-calibration variable
- Run a low regret least squares algorithm
- Make decision based on this forecast
## What bets to place?

<table>
<thead>
<tr>
<th>Bet Type</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>convex experts</td>
<td>([\hat{a}_t - a^*]_i)</td>
</tr>
<tr>
<td>internal regret</td>
<td>(e_{a^*} - e_{\hat{a}_t})</td>
</tr>
<tr>
<td>bandits</td>
<td>((e_a - e_b)I_{\hat{a}_t=b})</td>
</tr>
<tr>
<td>contextual</td>
<td>( \frac{I_{a_t=a}}{P(a_t=a)} - \frac{I_{a_t=\hat{a}_t}}{P(a_t=\hat{a}_t)} )</td>
</tr>
<tr>
<td>continuous</td>
<td>( (a_t - Mx_t)^2 )</td>
</tr>
<tr>
<td>LQR</td>
<td>( (a_t - \sum_{i=1}^{\log T} M_i x_{t-i})^2 )</td>
</tr>
<tr>
<td>reinforcement Learning</td>
<td>TD learn</td>
</tr>
<tr>
<td>Category</td>
<td>Bet</td>
</tr>
<tr>
<td>----------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>convex experts</td>
<td>$[\hat{a}_t - a^*]_i$</td>
</tr>
<tr>
<td>internal regret</td>
<td>$e_{a^*} - e_{\hat{a}_t}$</td>
</tr>
<tr>
<td>bandits</td>
<td>$(e_a - e_b)I_{\hat{a}_t = b}$</td>
</tr>
<tr>
<td>contextual</td>
<td>$\frac{I_{a_t = a}}{P(a_t = a)} - \frac{I_{a_t = \hat{a}_t}}{P(a_t = \hat{a}_t)}$</td>
</tr>
<tr>
<td>continuous</td>
<td>$X_t \times \left( \frac{I_{a_t = a}}{P(a_t = a)} - \frac{I_{a_t = \hat{a}_t}}{P(a_t = \hat{a}_t)} \right)$</td>
</tr>
<tr>
<td>LQR</td>
<td>$(a_t - Mx_t)^2$</td>
</tr>
<tr>
<td>reinforcement Learning</td>
<td>$(a_t - \sum_{i=1}^{\log T} M_i x_{t-i})^2$</td>
</tr>
<tr>
<td></td>
<td>TD learn</td>
</tr>
</tbody>
</table>
Theorem (Dicker 2019)

Least squares plus the calibration variable generates an estimate of the RL value function with low Macau.

Theorem (Dicker 2019)

A tweaked version of TD learning with $1/\sqrt{T}$ rates generates an estimate of the RL value function with low Macau.
Theorem (Dicker 2019)

*Least squares plus the calibration variable generates an estimate of the RL value function with low Macau.*


Theorem (Dicker 2019)

*A tweaked version of TD learning with $1/\sqrt{T}$ rates generates an estimate of the RL value function with low Macau.*

Proof: Similar to Dicker and F. 2018.
Conclusions

Take Aways

\[ \text{crazy-Calibration} + \text{low-regret} \iff \text{low-macau} \implies \text{good decisions} \]
Conclusions

Take Aways

crazy-Calibration + low-regret ⇐⇒ low-macau ⇒ good decisions

Thanks!
Proofs by example:

- low Regret $\nRightarrow$ low Macau
- low Regret $\nLeftarrow$ low Macau

Bets:

- Experts
- No Internal Regret
- Bandits, (scalar version), (exploration).
- Contextual Bandits
- Continuous action contextual Bandits
- Convex optimization, (one point), ($1/T$ with smooth)
- Reinforcement Learning
- LQR
Proofs: Regret ⇐⇒ macau
<table>
<thead>
<tr>
<th>$t$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>...</th>
<th>T-1</th>
<th>T</th>
<th>T+1</th>
<th>T+2</th>
<th>T+3</th>
<th>...</th>
<th>3T</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_t$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>$X_t$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>...</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>$\hat{Y}_t$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td>$\frac{1}{T}$</td>
<td>$\frac{2}{T+1}$</td>
<td>$\frac{3}{T+2}$</td>
<td>...</td>
<td>$\frac{2}{3}$</td>
</tr>
</tbody>
</table>

no regret $\iff$ not falsified
On-line least squares suffers no-regret:

- $\beta_t$ minimizes $\sum_{i=1}^{t} (Y_i - \beta \cdot X_t)^2$
- $\hat{Y}_t = \beta_{t-1} \cdot X_t$
- Total error: $\sum (Y_t - \hat{Y}_t)^2 = \min_\beta \sum (Y_t - \beta X_t)^2 + 4/9$
- In general, on-line least squares has $\log(T)$ total regret
- In this case, it actually wins by about $O(1)$. 

<table>
<thead>
<tr>
<th>$t$</th>
<th>$1$</th>
<th>$2$</th>
<th>$3$</th>
<th>$4$</th>
<th>$\cdots$</th>
<th>$T-1$</th>
<th>$T$</th>
<th>$T+1$</th>
<th>$T+2$</th>
<th>$T+3$</th>
<th>$\cdots$</th>
<th>$3T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_t$</td>
<td>$0$</td>
<td>$0$</td>
<td>$0$</td>
<td>$0$</td>
<td>$\cdots$</td>
<td>$0$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$\cdots$</td>
<td>$1$</td>
</tr>
<tr>
<td>$X_t$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$\cdots$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$\cdots$</td>
<td>$1$</td>
</tr>
<tr>
<td>$\hat{Y}_t$</td>
<td>$0$</td>
<td>$0$</td>
<td>$0$</td>
<td>$0$</td>
<td>$\cdots$</td>
<td>$0$</td>
<td>$0$</td>
<td>$\frac{1}{T}$</td>
<td>$\frac{2}{T+1}$</td>
<td>$\frac{3}{T+2}$</td>
<td>$\cdots$</td>
<td>$\frac{2}{3}$</td>
</tr>
</tbody>
</table>
\begin{align*}
\text{How about a bet?}
\end{align*}
No regret $\iff$ not falsified

<table>
<thead>
<tr>
<th>$t$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>$\cdots$</th>
<th>T-1</th>
<th>T</th>
<th>T+1</th>
<th>T+2</th>
<th>T+3</th>
<th>$\cdots$</th>
<th>3T</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_t$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
<td>1</td>
</tr>
<tr>
<td>$X_t$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
<td>1</td>
</tr>
<tr>
<td>$\hat{Y}_t$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
<td>0</td>
<td>1/T</td>
<td>2/(T+1)</td>
<td>3/(T+2)</td>
<td>$\cdots$</td>
<td>2/3</td>
</tr>
</tbody>
</table>

How about a bet?
No regret $\iff$ not falsified

<table>
<thead>
<tr>
<th>$t$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>$\cdots$</th>
<th>T-1</th>
<th>T</th>
<th>T+1</th>
<th>T+2</th>
<th>T+3</th>
<th>$\cdots$</th>
<th>3T</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_t$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
<td>1</td>
</tr>
<tr>
<td>$X_t$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\cdots$</td>
<td>1</td>
</tr>
<tr>
<td>$\hat{Y}_t$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
<td>0</td>
<td>$\frac{1}{T}$</td>
<td>$\frac{2}{T+1}$</td>
<td>$\frac{3}{T+2}$</td>
<td>$\cdots$</td>
<td>$\frac{2}{3}$</td>
</tr>
</tbody>
</table>

How about a bet?

- $Y_t > \hat{Y}_t$, so that is a safe bet!
- Construct this bet only using $X_t$

$$\sum_{i=1}^{T} X_t(Y - \hat{Y}_t) \approx T \frac{\log_e(3)}{2}$$

- Betting loses $\Omega(T)$
No regret $\iff$ not falsified

<table>
<thead>
<tr>
<th>$t$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>...</th>
<th>T-1</th>
<th>T</th>
<th>T+1</th>
<th>T+2</th>
<th>T+3</th>
<th>...</th>
<th>3T</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_t$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>$X_t$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>...</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>$\hat{Y}_t$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td>$\frac{1}{T}$</td>
<td>$\frac{2}{T+1}$</td>
<td>$\frac{3}{T+2}$</td>
<td>...</td>
<td>$\frac{2}{3}$</td>
</tr>
</tbody>
</table>

- Regret is $O(1)$
- Macau is $\frac{T}{2}$
- So: low regret $\iff$ low macau
Not falsified $\iff$ no regret

<table>
<thead>
<tr>
<th>$t$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>...</th>
<th>T</th>
<th>T+1</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_t$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>...</td>
<td>0</td>
<td>1</td>
<td>...</td>
</tr>
<tr>
<td>$X_t$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>...</td>
<td>1</td>
<td>1</td>
<td>...</td>
</tr>
<tr>
<td>$\hat{Y}_t$</td>
<td>.6</td>
<td>.4</td>
<td>.6</td>
<td>.4</td>
<td>...</td>
<td>.6</td>
<td>.4</td>
<td>...</td>
</tr>
</tbody>
</table>
Betting

- No bet based on $X_t$ will win anything
- In other words,

$$\max_{\alpha} \sum_{i=1}^{T} \alpha \cdot X_t (Y - \hat{Y}_t) = 0$$

- This forecast is not falsified using linear functions of $X_t$
Not falsified $\iff$ no regret

<table>
<thead>
<tr>
<th>$t$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>...</th>
<th>T</th>
<th>T+1</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_t$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>...</td>
<td>0</td>
<td>1</td>
<td>...</td>
</tr>
<tr>
<td>$X_t$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>...</td>
<td>1</td>
<td>1</td>
<td>...</td>
</tr>
<tr>
<td>$\hat{Y}_t$</td>
<td>.6</td>
<td>.4</td>
<td>.6</td>
<td>.4</td>
<td>...</td>
<td>.6</td>
<td>.4</td>
<td>...</td>
</tr>
</tbody>
</table>

But, a better forecast exists

- $\sum (Y_t - \hat{Y}_t)^2 = .36T$
- $\min_\beta (Y_t - \beta X_t)^2 = .25T$
- Regret is $.11T$
- So, regret is $\Omega(T)$
Not falsified $\iff$ no regret

<table>
<thead>
<tr>
<th>$t$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>...</th>
<th>T</th>
<th>T+1</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_t$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>...</td>
<td>0</td>
<td>1</td>
<td>...</td>
</tr>
<tr>
<td>$X_t$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>...</td>
<td>1</td>
<td>1</td>
<td>...</td>
</tr>
<tr>
<td>$\hat{Y}_t$</td>
<td>.6</td>
<td>.4</td>
<td>.6</td>
<td>.4</td>
<td>...</td>
<td>.6</td>
<td>.4</td>
<td>...</td>
</tr>
</tbody>
</table>

- Macau is zero
- Regret is $T/9$
- So: low macau $\iff$ low regret
In the convex optimization problem, we observe a sequence of convex functions $c_t(\cdot)$. Our goal is to figure out an action $\hat{x}_t^*$ to take at each point in time $t$ to minimize $\sum_t c_t(\hat{x}_t^*)$.

- Forecast: Gradient of $c_t$ at each point in time $t$
  \( g_t(x) \equiv \nabla c_t(x) \)
- Strategy: Pick a $\hat{x}_t^*$ such that $\hat{g}_t(\hat{x}_t^*) = 0$.
- Worry: “The real optimum $x^*$ would generate better performance.”

Macau bets: $[x^* - \hat{x}_t^*]_i$ bet against $[g_t]_i - [\hat{g}_t]_i$

\[
\text{Macau}_i = \sum_{t=1}^{T} [x^* - \hat{x}_t^*]_i ([g_t]_i - [\hat{g}_t]_i)
\]

Bet: $[x^* - \hat{x}_t^*]_i$
In the convex optimization problem, we observe a sequence of convex functions $c_t(\cdot)$. Our goal is to figure out an action $\hat{x}^*_t$ to take at each point in time $t$ to minimize $\sum_t c_t(\hat{x}^*_t)$.

- Forecast: $c_t(x)$ at points near $\hat{x}^*_t$, for example $x_t - \hat{x}^*_t \sim N(0, \sigma^2 I)$
- Strategy: Pick a $\hat{x}^*_t$ to minimize $\hat{c}(\cdot)$
- Worry: “The real optimum $x^*$ would generate better performance.”
- Macau bets: $(x^* - \hat{x}^*_t) \cdot (x_t - \hat{x}^*_t)$

$$\text{Macau} = \sum_{t=1}^{T} (x^* - \hat{x}^*_t) \cdot (x_t - \hat{x}^*_t) c(x)$$

Bet: $[x^* - \hat{x}^*_t]_i$
Bet: Optimizing continuous convex functions (with gradient)

Also assume each $c_t$ is smooth, say $c_t \in C^2$. We’ll keep all else the same.

- We can use the macau to look at bets for how for $\hat{\beta}$ is from the best after the fact $\beta$
- Thus we know the optimum point is close to the best hind sight decision point (say $1/\sqrt{T}$ accuracy)
- This means the error in payoff space is $1/T$
- So it doesn’t require a new algorithm or even new features
In the experts problem, we observe the payoff of $k$ different experts. Our goal is to generate as much value as the best expert.

- Forecast: one value for each arm ($Y_t \in \mathbb{R}^k$, so $\hat{Y}_t \in \mathbb{R}^k$ also)
- Strategy: Pick arm with highest forecast ($\hat{a}_t = \arg \max_i [\hat{Y}_t]_i$)
- Worry: “Always playing arm $b$ would generate more”
- Macau bet: $e_b = [0, 0, 0, \ldots, 1, \ldots, 0]'$

$$\text{Macau} = \max_{b \in \{1, \ldots, k\}} \sum_t (e_b - e_{\hat{a}_t}) \cdot (Y_t - \hat{Y}_t)$$

Bet: $e_b - e_{\hat{a}_t}$
Bet: No Internal Regret

In the no-internal regret problem, we observe the payoff of $k$ different experts. Our goal is to avoid feeling regret about possibly switching one of our actions to some other action.

- Forecast: one value for each expert ($Y_t \in \mathbb{R}^k$, so $\hat{Y}_t \in \mathbb{R}^k$ also)
- Strategy: Pick arm with highest forecast ($\hat{a}_t = \arg \max_i [\hat{Y}_t]_i$)
- Worry: “Playing $c$ when we previously played $b$ would have been better ($R^{c \rightarrow b} > 0$).”
- Macau bet:

$$ (I_{\hat{a}_t = c} (e_b - e_c)) \cdot (Y_t - \hat{Y}_t) $$

Bet on $c \rightarrow b$: $I_{\hat{a}_t = c} (e_b - e_c)$
The rest isn’t done yet!
We only see outcomes on the one of \( k \) arms we pull.

- Forecast: Each arms payoff: \( [Y_t]_i = \frac{r_t 1_{a_t=i}}{p(a_t=i)} \), so \( \hat{Y}_t \in \mathbb{R}^k \).
- Strategy: Pick arm with highest forecast \( \hat{a}_t = \arg \max_i [\hat{Y}_t]_i \) with some exploration also.
- Worry: Always playing \( b \) might have been better.
- Macau bet:
  \[
  (e_b - e_{\hat{a}_t}) \cdot (Y_t - \hat{Y}_t)
  \]

Bet on \( b \): \( (e_b - e_{\hat{a}_t}) \)
Play $a_t \in \{1, \ldots, k\}$ and only see its outcome.

- Forecast: the arm actually played: $Y_t = \frac{r_t(a_t)}{p_t(a_t)}$, so $\hat{Y}_t(a_t) \in \mathbb{R}$.

- Strategy: Pick arm with highest forecast ($\hat{a}_t = \arg \max_i \hat{Y}_t(i)$) with some exploration also.

- Worry: Always playing $b$ might have been better.

- Macau bet:

$$
\left( \frac{l_{a_t=b}}{p_t(b)} - \frac{l_{a_t=\hat{a}_t}}{p_t(\hat{a}_t)} \right) (Y_t - \hat{Y}_t)
$$

Bet on $b$: $\frac{l_{a_t=b}}{p_t(b)} - \frac{l_{a_t=\hat{a}_t}}{p_t(\hat{a}_t)}$
Bandits exploration

- Macau keeps the mean correct
- We would also high probability statements
- So, we need $p_t(b)$ to not be too small
  - Easy math: $p_t(b) \geq t^{-1/3}$, but not optimal rates of convergence
  - Giving up a log: $p_t(b) \geq t^{-1/2}$. But, as $\hat{Y}_t(b)$ gets closer to $\hat{Y}_t(\hat{a}_t)$ we sample more often. On a log scale, this means we need $k \log(T)$ features.
  - Note: the fixed point solution will generate some randomization above and beyond that given by the lower bounds
- Similar behavior to UCB, but a different philosophy to justify it.
Bet: Contextual Bandits (vector version)

First we observe $X_t \in \mathbb{R}^d$, then we play an arm $a_t$ and observe its outcome (vector version: $[Y_t]_i = \frac{r_{tI_{a_t=i}}}{p(a_t=i)}$):

- Forecast: $\hat{Y}_t = X_t\beta_{t-1}$, with $\beta \in \mathbb{R}^{d \times k}$ $\hat{Y}_t \in \mathbb{R}^k$.
- Strategy: Pick arm with highest forecast ($\hat{a}_t = \text{arg max}_i [\hat{Y}_t]_i$).
- Worry: Using some other $\beta^*$ might be better.
- Naive Macau bet ($\hat{a}_t \rightarrow b$):

$$ (I_{X_t(\beta^*_b - \beta^*_\hat{a}_t)}>0 - e_{\hat{a}_t}) \cdot (Y_t - \hat{Y}_t) $$

These are hard to put in a linear space. But, given the low dimension ($VC=d+2$) hope spring eternal.

Bet on $b$: $\quad (e_b - e_{\hat{a}_t})$
Bet: Continuous action for contextual Bandits

First we observe $X_t \in \mathbb{R}^d$, then we play an action $a_t \in \mathcal{A} \subset \mathbb{R}^k$ and observe its outcome. (We’ll actually penalize $a$ quadratically and hence avoid the set $\mathcal{A}$.)

- **Forecast**: $\hat{Y}_t(a) = X_t^\top \beta_{t-1} a - a^\top a/2$, with $\beta \in \mathbb{R}^{d \times k}$ and $\hat{Y}_t(a) \in \mathbb{R}^k$.

- **Strategy**: Pick “best” action: $\hat{a}_t = \arg \max_{a \in \mathcal{A}} \hat{Y}_t(a) = X_t^\top \hat{\beta}_{t-1}$.

- **Worry**: Using some other $\beta^*$ might be better.

- **Naive Macau bet** ($\hat{a}_t \rightarrow (1 - \epsilon)\hat{a}_t + \epsilon X_t^\top \beta^*$):

$$
(X_t^\top \beta^* - X_t^\top \hat{\beta}_t^*) \cdot (a_t - \hat{a}_t)(Y_t(a_t) - \hat{Y}_t(a_t))
$$

Bet in direction $X_t^\top \beta^*$: (fillin)
The RL value function:

$$V_t^* = \max_{\pi} \mathbb{E} \left( \sum_{i=t}^{\infty} \gamma^{i-t} r_i(a_{i}^{\pi}) \mid \mathcal{F}_t \right)$$

($\gamma$ is discount rate.) Recursively:

$$V_t^* = \mathbb{E} (r_t(a) + \gamma V_{t+1}^* \mid \mathcal{F}_t)$$
The RL value function:

\[ V_t^* = \max_{\pi} E \left( \sum_{i=t}^{\infty} \gamma^{i-t} r_i(a^\pi_i) \bigg| F_t \right) \]

(\(\gamma\) is discount rate.) Recursively:

\[ V_t^* = E \left( r_t(a) + \gamma V_{t+1}^* \big| F_t \right) \]

\(V^*\) is a Y-variable and an X-variable!