Macau: Why talk about regret when betting is so much more fun?

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.)

Solution: Falsifiable is better. A falsifiable forecast can be bet against. The amount lost to the adversary is its 
macau 

Take Aways
crazy-Calibration + low-regret = low-macau = good decisions
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.
- A falsifiable forecast can be bet against
- The amount lost to the adversary is its macau.
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.
- A falsifiable forecast can be bet against
- The amount lost to the adversary is its macau.

Take Aways

\[ \text{crazy-Calibration} + \text{low-regret} \implies \text{low-macau} \implies \text{good decisions} \]
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
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:
    \[
    \text{Winnings} = E \left( B (Y - \hat{Y}) \right)
    \]
    - $(Y - \hat{Y})$ is a “fair” bet
    - $B$ is amount 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:
    \[
    \text{Winnings} = E \left( B (Y - \hat{Y}) \right)
    \]
  - \( (Y - \hat{Y}) \) 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:
    \[
    \text{Winnings} = E \left( B \left( Y - \hat{Y} \right) \right)
    \]
  - \( Y - \hat{Y} \) is a “fair” bet
  - \( B \) is amount bet
- How to avoid being proven wrong by:
  \[
  \text{Macau} \equiv \max_{|B| \leq 1} E \left( B \left( Y - \hat{Y} \right) \right)
  \]
  (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:
    \[
    \text{Winnings} = E \left( B (Y - \hat{Y}) \right)
    \]
  - $(Y - \hat{Y})$ 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)}
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>
</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 \( t \)
Crazy calibration variable

\[ Y \]

\[ X_1 \quad X_2 \quad X_3 \quad X_4 \]

| \( Y_1 \) | \( X_{11} \) | \( X_{12} \) | \( X_{13} \) | \( X_{14} \) |
| \( Y_2 \) | \( X_{21} \) | \( X_{22} \) | \( X_{23} \) | \( X_{24} \) |
| \( Y_3 \) | \( X_{31} \) | \( X_{32} \) | \( X_{33} \) | \( X_{34} \) |
| \( Y_4 \) | \( X_{41} \) | \( X_{42} \) | \( X_{43} \) | \( X_{44} \) |
| \( \vdots \) | \( \vdots \) | \( \vdots \) | \( \vdots \) | \( \vdots \) |
| \( Y_t \) | \( X_{t1} \) | \( X_{t2} \) | \( X_{t3} \) | \( X_{t4} \) |

\[ \hat{\beta}_t = \arg\min_{\beta} \sum_{i=1}^{t} (Y_i - \beta'X_i)^2 \]

We can fit \( \hat{\beta}_t \) on everything up to time \( t \)
From a new $X_{t+1}$'s 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

<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>$\vdots$</td>
<td>$\vdots$</td>
<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>
<td>$\hat{\beta}_{t-1}$</td>
<td>$\hat{Y}<em>t = \hat{\beta}</em>{t-1}' X_t$</td>
</tr>
</tbody>
</table>

**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></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>$Y_t$</td>
<td>$X_{t1}$</td>
<td>$X_{t2}$</td>
<td>$X_{t3}$</td>
<td>$X_{t4}$</td>
</tr>
</tbody>
</table>

$\hat{\beta} \rightarrow \hat{Y}$

$0 \rightarrow \hat{Y}_1 = 0$

$\hat{\beta}_1 \rightarrow \hat{Y}_2 = \hat{\beta}_1' X_2$

$\hat{\beta}_2 \rightarrow \hat{Y}_3 = \hat{\beta}_2' X_3$

$\hat{\beta}_3 \rightarrow \hat{Y}_4 = \hat{\beta}_3' X_4$

$\hat{\beta}_{t-1} \rightarrow \hat{Y}_t = \hat{\beta}_{t-1}' X_t$

**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

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

$$\hat{Y} = \hat{\beta}_{t-1} X_t$$

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>$\hat{Y}$</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>$\hat{\beta}$</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>$\hat{Y}_2 = \hat{\beta}_1^t X_2$</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>$\hat{Y}_3 = \hat{\beta}_2^t X_3$</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>$\hat{Y}_4 = \hat{\beta}_3^t X_4$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<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>$\hat{Y}_t$</td>
<td>$X_{t4}$</td>
<td>$\hat{\beta}_{t-1}$</td>
<td>$\hat{Y}<em>t = \hat{\beta}</em>{t-1}^t X_t$</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Y</th>
<th>X₁</th>
<th>X₂</th>
<th>X₃</th>
<th>X₄</th>
<th>( \hat{\beta} )</th>
<th>( \hat{Y} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y₁</td>
<td>X₁₁</td>
<td>X₁₂</td>
<td>( \hat{Y}_₁ )</td>
<td>X₁₄</td>
<td>0</td>
<td>( \hat{Y}_₁ = 0 )</td>
</tr>
<tr>
<td>Y₂</td>
<td>X₂₁</td>
<td>X₂₂</td>
<td>( \hat{Y}_₂ )</td>
<td>X₂₄</td>
<td>( \hat{\beta}_₁ )</td>
<td>( \hat{Y}_₂ = \hat{\beta}_₁ X₂ )</td>
</tr>
<tr>
<td>Y₃</td>
<td>X₃₁</td>
<td>X₃₂</td>
<td>( \hat{Y}_₃ )</td>
<td>X₃₄</td>
<td>( \hat{\beta}_₂ )</td>
<td>( \hat{Y}_₃ = \hat{\beta}_₂ X₃ )</td>
</tr>
<tr>
<td>Y₄</td>
<td>X₄₁</td>
<td>X₄₂</td>
<td>( \hat{Y}_₄ )</td>
<td>X₄₄</td>
<td>( \hat{\beta}_₃ )</td>
<td>( \hat{Y}_₄ = \hat{\beta}_₃ X₄ )</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Yₜ</td>
<td>Xₜ₁</td>
<td>Xₜ₂</td>
<td>( \hat{Y}_ₜ )</td>
<td>Xₜ₄</td>
<td>( \hat{\beta}_{t-1} )</td>
<td>( \hat{Y}<em>ₜ = \hat{\beta}</em>{t-1} Xₜ )</td>
</tr>
</tbody>
</table>

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

Adding the crazy calibration variable generates low macau:

\[
\sum_{t=1}^{T} X_t(Y_t - \hat{Y}_t) = O(\sqrt{T \log(T)})
\]
Macau is a “normal equation”

\[ E(Y|X) \]

<table>
<thead>
<tr>
<th>Least squares</th>
<th>Normal equations</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \min_\beta \sum (Y_i - \beta \cdot X_i)^2 )</td>
<td>( \sum X_i (Y_i - \beta \cdot X_i) = 0 )</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 is a “normal equation”

\[ E(Y|X) \]

**Statistics**

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

- **Normal equations**
  \[ \min_{\beta} \max_{\alpha} \sum \alpha \cdot X_i \ (Y_i - \beta \cdot X_i) \]
Macau is a “normal equation”

| Statistics | E(Y|X) | Least squares | Normal equations |
|------------|-------|---------------|------------------|
|            | 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) | |
| Probability| min_{f} E \left( (Y - f(X))^2 \right) | (\forall g) E(g(X) (Y - f(X))) = 0 | |

The normal equation is the same as:

$$\max_{g} E \left( g(X)(Y - f(X)) \right) = 0$$

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

$$\min_{f} \max_{g} E \left( g(X)(Y - f(X)) \right) = 0$$
Macau is a “normal equation”

| $E(Y|X)$ | Least squares | Normal equations |
|----------|---------------|------------------|
| **Statistics** | $\min_\beta \sum (Y_i - \beta \cdot X_i)^2$ | $\min_\beta \max_\alpha \sum \alpha \cdot X_i \ (Y_i - \beta \cdot X_i)$ |
| **Probability** | $\min_f E \left( (Y - f(X))^2 \right)$ | $\min_f \max_g E \left( g(X) \ (Y - f(X)) \right)$ |
Macau is a “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>$\min_f E \left( (Y - f(X))^2 \right)$</td>
<td>$\min_f \max_g E \left( g(X) \ (Y - f(X)) \right)$</td>
</tr>
<tr>
<td><strong>online</strong></td>
<td>low regret</td>
<td>low macau</td>
</tr>
</tbody>
</table>

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

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

| online | low regret | low macau |

$$Macau \equiv \max_{\alpha:|\alpha| \leq 1} \sum_{t=1}^{T} \alpha \cdot X_t \left( Y_t - \hat{Y}_t \right)$$
Macau is a “normal equation”

| E(Y|X)       | Least squares                                                                 | Normal equations                                                                 |
|--------------|-------------------------------------------------------------------------------|----------------------------------------------------------------------------------|
| Statistics   | \( \min_\beta \sum (Y_i - \beta \cdot X_i)^2 \)                            | \( \min_\beta \max_\alpha \sum \alpha \cdot X_i \ (Y_i - \beta \cdot X_i) \) |
| Probability  | \( \min_f E \left( (Y - f(X))^2 \right) \)                                | \( \min_f \max_g E\left( g(X) \ (Y - f(X)) \right) \)                        |
| online       | low regret                                                                    | low macau                                                                       |

- probability: Least squares \( \iff \) normal equations
- statistics: Least squares \( \iff \) normal equations
Macau is a “normal equation”

| E(Y|X) | Least squares | Normal equations |
|--------|---------------|------------------|
| Statistics | \( \min_\beta \sum (Y_i - \beta \cdot X_i)^2 \) | \( \min_\beta \max_\alpha \sum \alpha \cdot X_i \ (Y_i - \beta \cdot X_i) \) |
| Probability | \( \min_f E \left( (Y - f(X))^2 \right) \) | \( \min_f \max_g E \left( g(X) \ (Y - f(X)) \right) \) |
| online | low regret | low macau |

Take Aways

on-line low regret \( \iff \) on-line low macau
No regret $\not\Rightarrow$ 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>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>$\hat{\gamma}_t$</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 $\not\Rightarrow$ 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>$\gamma_t$</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>
</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>$\hat{\gamma}_t$</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>
</tr>
</tbody>
</table>

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

How about a bet?

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>0</th>
<th>0.02</th>
<th>0.04</th>
<th>0.06</th>
<th>0.08</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>time</td>
<td>0</td>
<td>50</td>
<td>100</td>
<td>150</td>
<td>200</td>
<td>300</td>
</tr>
</tbody>
</table>

No regret $\iff$ not falsified

$\not\Rightarrow$ low regret
No regret $\nRightarrow$ not falsified

$$
\begin{array}{cccccccc}
  t & 1 & 2 & 3 & 4 & \cdots & T-1 & T & T+1 & T+2 & T+3 & \cdots & 3T \\
  Y_t & 0 & 0 & 0 & \cdots & 0 & 1 & 1 & 1 & 1 & \cdots & 1 \\
  X_t & 1 & 1 & 1 & \cdots & 1 & 1 & 1 & 1 & 1 & \cdots & 1 \\
  \hat{Y}_t & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & \frac{2}{T+1} & \frac{3}{T+2} & \cdots & \frac{2}{3} \\
\end{array}
$$

Not falsified $\nRightarrow$ no regret

$$
\begin{array}{cccccccc}
  t & 1 & 2 & 3 & 4 & \cdots & T & T+1 & \cdots \\
  Y_t & 0 & 1 & 0 & 1 & \cdots & 0 & 1 & \cdots \\
  X_t & 1 & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\
  \hat{Y}_t & .6 & .4 & .6 & .4 & \cdots & .6 & .4 & \cdots \\
\end{array}
$$

- Macau is zero
- Regret is $T/9$
- So: low macau $\nRightarrow$ 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:

- $a$ = action taken $\in \mathbb{R}^k$ (e.g., inventory levels)
- $X_t$ = Context at time $t$
- $a_t^*$ = best action at time $t$
- $r_t(a)$ = Reward at time $t$ playing $a$
- $V_t^*$ = $\max_a E(r_t(a)|X_t) = E(r_t(a^*)|X_t)$
- $q_t(a) \leq E(r_t(a)|X_t) \leq \bar{q}_t(a)$

What are good falsifiable claims about $a^*$?
Contextual Bandits

Some notation:

\[ a \in \mathbb{R}^k \text{(eg inventory levels)} \]

\[ X_t = \text{Context at time } t \]

\[ a^*_t = \text{best action at time } t \]

\[ r_t(a) = \text{Reward at time } t \text{ playing } a \]

\[ V^*_t = \max_a E(r_t(a)|X_t) = E(r_t(a^*)|X_t) \]

\[ q_t(a) \leq E(r_t(a)|X_t) \leq \overline{q}_t(a) \]

Too precise:

- “\( q_t(a) = \overline{q}_t(a) \)”
Contextual Bandits

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 \( a^*_t \).”
Contextual Bandits

Some notation:

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

Just right:

\[
\begin{align*}
    \bar{q}_t(a) & = V_t^* - q||a - a^*_t||^2 \\
    \bar{q}_t(a) - q_t(a) & = \Delta||a - a^*_t||^2
\end{align*}
\]
Why is low macau useful?

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

- Supposed each \( c_t(\cdot) \) is convex
- We could use \( \nabla c_t(\cdot) \) for SGD
- Say, on-line convex optimization:

\[
\text{regret} = \sum_{t=1}^{T} (c_t(\hat{a}_t) - c_t(a^*))
\]

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

$$C(a) = \sum_{t=1}^{T} c_t(a) \quad a^* \equiv \arg \min_{\hat{a}} C(\hat{a})$$

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:

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

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

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

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

\[ \text{(zero @ } \hat{a}_t) \]

(macau!)
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:

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

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

\[
\text{regret} \leq \text{macau}
\]
without crazy-calibration variable
Using the crazy-calibration variable

Predicted

![Graph showing the relationship between predicted and actual values]
Theorem ( \(\rightarrow\) F. and Kakade 2008, \(\leftarrow\) 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 (\(\iff\) 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 (an Amazon scholar!) and D. Foster seem to have a proof of something very close to this for IID.)
Theorem (\(\iff\) 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 \(M = o(T)\).

Proof sketch: Consider the forecasts \(w\hat{Y}_t + (1 - w)\beta \cdot X_t\) for the best \(\beta\). Let \(Q(w)\) be the total quadratic error of this forecast. Best \(\beta\) says, the minimum is at \(Q(1)\). The following three are equivalent:

- \(Q(0) \leq Q(1)\) (No regret condition)
- \(Q'(0)\) is zero. (No macau condition)
- \(Q(w) = 0\) for all \(w\).
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>Category</th>
<th>Bet</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>
</tbody>
</table>

**reinforcement Learning**

TD learn
<table>
<thead>
<tr>
<th>What bets to place?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bet</strong></td>
</tr>
<tr>
<td>convex</td>
</tr>
<tr>
<td>experts</td>
</tr>
<tr>
<td>internal regret</td>
</tr>
<tr>
<td>bandits</td>
</tr>
<tr>
<td>contextual</td>
</tr>
<tr>
<td>continuous</td>
</tr>
<tr>
<td>LQR</td>
</tr>
<tr>
<td>reinforcement Learning</td>
</tr>
<tr>
<td>Theorem (Dicker 2019)</td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td>Least squares plus the calibration variable generates an estimate of the RL value function with low Macau.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Theorem (Dicker 2019)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A tweaked version of TD learning with $1/\sqrt{T}$ rates generates an estimate of the RL value function with low Macau.</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.*

Proof: Similar to Dicker and F. 2018.
Conclusions

Take Aways

crazy-Calibration + low-regret $\iff$ low-macau $\implies$ good decisions
Conclusions

Take Aways

crazy-Calibration + low-regret $\iff$ low-macau $\implies$ good decisions

Thanks!
Proofs by example:

- low Regret $\iff$ low Macau
- low Regret $\niff$ 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
No regret $\iff$ not falsified

<table>
<thead>
<tr>
<th>$t$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>$T-1$</th>
<th>$T$</th>
<th>$T+1$</th>
<th>$T+2$</th>
<th>$T+3$</th>
<th>$\ldots$</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>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\ldots$</td>
<td>1</td>
</tr>
<tr>
<td>$X_t$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\ldots$</td>
<td>1</td>
</tr>
<tr>
<td>$\hat{Y}_t$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</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>$\ldots$</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 $O(1)$. 
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?
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?

no regret $\iff$ not falsified
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 2 3 4 ··· T-1 T T+1 T+2 T+3 ··· 3T</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_t$</td>
<td>0 0 0 0 ··· 0 1 1 1 1 1 ··· 1</td>
</tr>
<tr>
<td>$X_t$</td>
<td>1 1 1 1 ··· 1 1 1 1 1 1 ··· 1</td>
</tr>
<tr>
<td>$\hat{Y}_t$</td>
<td>0 0 0 0 ··· 0 0 $\frac{1}{T}$ $\frac{2}{T+1}$ $\frac{3}{T+2}$ ··· $\frac{2}{3}$</td>
</tr>
</tbody>
</table>

- Regret is $O(1)$
- Macau is $T/2$
- 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>
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>.6</td>
<td>.4</td>
<td>.6</td>
<td>.4</td>
<td>$\cdots$</td>
<td>.6</td>
<td>.4</td>
<td>$\cdots$</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

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

| $t$ | 1  | 2  | 3  | 4  | ... | T  | T+1 | ...
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></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:** $(x^* - \hat{x}_t^*) \cdot (x_t - \hat{x}_t^*)$

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

Bet: $[x^* - \hat{x}_t^*]_i$
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: $e_b - e_{\hat{a}_t}$
{dpf: I’m not sure this actually works. Needs work!}

Will least squares do well enough? Consider two arms

- First arm gets a known sequence of payoffs (so forecast is trivial, $r_t = V_t$)
- Second arm forecast has no-regret but lots of Macau:

\[ Y \]

- We might play first arm—but we could bet the second arm would be better and win!
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:
  \[
  (l_{\hat{a}_t = c}(e_b - e_c)) \cdot (Y_t - \hat{Y}_t)
  \]

Bet on $c \rightarrow b$: $l_{\hat{a}_t = c}(e_b - e_c)$
We only see outcomes on the one of $k$ arms we pull.

- **Forecast:** Each arms payoff: $[Y_t]_i = \frac{r_t/l_{a_t=i}}{p(a_t=i)}$, 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 ($\hat{a}_t = \text{arg max}_i \hat{Y}_t(i)$) 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 $b$: $\frac{l_{a_t=b}}{p_t(b)} - \frac{l_{a_t=\hat{a}_t}}{p_t(\hat{a}_t)}$
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: $[Y_t]_i = \frac{r_t I_{a_t=i}}{p(a_t=i)}$):

- **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 ($\hat{a}_t = \arg \max_i [\hat{Y}_t]_i$).

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

- **Naive Macau bet ($\hat{a}_t \rightarrow b$):**

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

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

Bet on $b$: $$(e_b - e_{\hat{a}_t})$$
First we observe $X_t \in \mathbb{R}^d$, then we play an action $a_t \in A \subset \mathbb{R}^k$ and observe its outcome. (We’ll actually penalize $a$ quadratically and hence avoid the set $A$.)

- **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: $\hat{a}_t = \arg\max_{a \in A} \hat{Y}_t(a) = X_t^\top \hat{\beta}_{t-1}$.

- **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 rest isn’t done yet!
The RL value function:

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

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

$$V_t^* = E \left( r_t(a) + \gamma V_{t+1}^* \middle| \mathcal{F}_t \right)$$
Reinforcement Learning

The RL value function:

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

(\( \gamma \) 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!
bets for LQR

Linear Quadratic Regulator: