LS88: Sports Analytics

Regression to the Mean

Regression to the Mean in a Nutshell

Given observed performance, future performance is expected to be closer to the expected value than observed performance

Pilot Performance

A famous example from Daniel Kahneman

- ➔ Fighter pilots were being graded on performance of their landings
- ➔ Pilot performs well ➡ pilot is praised ➡ pilot performance declines
- ➔ Pilot performs poorly ➡ pilot is punished ➡ pilot performance increases
- ➔ Instructor then thinks: punish bad performance to get improvement

In reality: performance is part skill, part luck/random

- ➔ Pilots would naturally *regress to their mean performance* after a poor flight because it was luck that day that made them look bad
- ➔ Punishment was ancillary and they were likely to see improvement in the pilot without intervention

Theoretical Statement

- ➔ Two measurements, M_1 and M_2 with mean μ

$$\mu \leq \text{Expected Value}[M_2 \ \ Given \ \ M_1 = c > \mu] < c$$

- ➔ Given the first measurement is above μ, we expect the second measurement to be closer to μ
- ➔ Similar statement for c < μ

Regression to the Mean in Sports

Rarely do we know μ exactly: for example, in sports

Early season performance compared to the rest of the season

- Brandon Belt hits .600 in first weeks of the season
  - We don't know μ but we know typical values for the league (it's been 77 years since Ted WIlliams hit .400 *for the season*)
  - We also know belt is a career .266 hitter: .600 is far above his usual rate
  - Expect performance after 2 weeks between .266 and .600

Regression to the Mean in Sports

- Klay Thompson is shooting 34% from 3 right now
  - We don't know μ but we know Klay is a career 42% 3pt shooter
  - Expect performance after today between 34% and 43%

Where should our expectation fall between the two choices? We'll get there

Outline

- ➔ Derive RTTM from regression
- ➔ Motivation from beliefs (Bayesian Inference)
- ➔ Applied to baseball batting averages
- ➔ Applied to projection systems: Marcels

Regression

Student Test Scores

Students have a certain talent level on a scale of 0-100

- ➔ The (unknown) average talent is 50
- ➔ The (unknown) standard deviation of talent is 5

Students are going to take two tests to assess talent

- ➔ The test is a noisy measurement of talent
- ➔ We assume the noise has expected value of 0 (doesn't bias the scores up or down from talent level)
- ➔ The (unknown) standard deviation of noise is 2

We want to predict the second score from the first test score

Student Test Scores

If we observe 1 score from 1 student, we just use that score to predict the next score

If we observe multiple scores from 1 student, we take the average

What if we observe 1 score from multiple students???

- ➔ Okay, so actually to build a model, we need to do this 2 test experiment to build the model
- ➔ THEN we can predict the second score from 1 score from another set of students

Regression Prediction

We gather pairs of test scores from *N* students

The regression model is:

$$\begin{aligned}\text{Test 2 Score} &= \alpha + \beta \times \text{Test 1 Score} \\ &= \alpha + \beta \times (\text{Test 1 Score} - \text{Test 1 Avg} + \text{Test 1 Avg}) \\ &= (\alpha + \beta \times \text{Test 1 Avg}) + \beta \times (\text{Test 1 Score} - \text{Test 1 Avg}) \\ &= \tilde{\alpha} + \beta \times (\text{Test 1 Score} - \text{Test 1 Avg})\end{aligned}$$

➔ If $\beta < 1$, *regression to the mean*

Let's see some examples

Regression Prediction

Formulas for regression coefficients

$$\text{Test 2 Score} = \hat{\alpha} + \hat{\beta} \times \text{Test 1 Score}$$

$$\hat{\alpha} = \text{Test 2 Avg} - \hat{\beta} \times \text{Test 1 Avg}$$

$$\hat{\beta} = \text{Corr}(\text{Test 1 Score}, \text{Test 2 Score}) \times \frac{\text{StdDev Test 2 Score}}{\text{StdDev Test 1 Score}}$$

Regression Prediction

More detailed regression equation

$$\begin{aligned}
\text{Test 2 Score} &= \text{Test 2 Avg} + \hat{\beta} \times (\text{Test 1 Score} - \text{Test 1 Avg}) \\
&= \text{Test 2 Avg} + \rho \times \text{StdDev Test 2 Score} \times \frac{\text{Test 1 Score} - \text{Test 1 Avg}}{\text{StdDev Test 1 Score}} \\
&= \text{Test 2 Avg} + \rho \times \text{StdDev Test 2 Score} \times \text{Test 1 Z-Score}
\end{aligned}$$

- ➔ Test 2 Avg controls level
- ➔ Std Dev of Test 2 is used to scale properly
- ➔ Correlation coefficient controls dependence on Test 1
  - ◆ If =1, use Test 1 score
  - ◆ If =0, Test 1 doesn't matter so use the average

Regression Prediction

More detailed regression equation

$$\begin{aligned} \text{Test 2 Score} &= \text{Test 2 Avg} + \hat{\beta} \times (\text{Test 1 Score} - \text{Test 1 Avg}) \\ &= \text{Test 2 Avg} + \rho \times \text{StdDev Test 2 Score} \times \frac{\text{Test 1 Score} - \text{Test 1 Avg}}{\text{StdDev Test 1 Score}} \\ &= \text{Test 2 Avg} + \rho \times \text{StdDev Test 2 Score} \times \text{Test 1 Z-Score} \end{aligned}$$

- ➔ Test 2 Avg controls level
- ➔ Std Dev of Test 2 is used to scale properly
- ➔ Correlation coefficient controls dependence on Test 1
    - ◆ If =1, use Test 1 score
    - ◆ If =0, Test 1 doesn't matter so use the average

Regression to the Mean

The general form:

$$\text{Regressed Estimate} = w \cdot \text{Observation} + (1 - w) \cdot \text{Mean}, \quad 0 \le w \le 1$$

Updating Beliefs

Coin Flipping

- ➔ We're going to flip a coin
- ➔ It has an unknown probability of heads equal to *p*
- ➔ As we observe tosses, we want to update our beliefs about *p*

What should our starting belief be? How do we update our beliefs?

This is all a non-technical primer on *Bayesian Inference*

- ➔ Bayesian Inference is neither superior or inferior, it's just a different approach

Initial Beliefs aka Prior Beliefs

We're going to start with an agnostic belief

- ➔ We have no idea what *p* could be other than it's between 0 and 1

We quantify our beliefs according to an number of expected heads (H) and expected tails (T)

- ➔ The larger either or both H and T are, the more firm we are in our beliefs
- ➔ Our agnostic belief is H = T = 1
- ➔ We'll eventually see the formula for producing our belief from H and T

$$\text{Best Guess } p = \frac{H}{H+T}$$

Updating Beliefs

We toss the coin and have updated counts for H and T

$$H_{new} = H_{old} + \text{Number of Heads tossed}, \quad T_{new} = T_{old} + \text{Number of Tails tossed}$$

Our best is guess is always the fraction of heads:

$$\begin{aligned}
\text{Updated Best Guess for } p &= \frac{H_{new}}{H_{new} + T_{new}} \\
&= \frac{H_{old} + \text{Number of Heads tossed}}{H_{old} + \text{Number of Heads tossed} + T_{old} + \text{Number of Tails tossed}} \\
&= \frac{H_{old} + \text{Number of Heads tossed}}{H_{old} + T_{old} + \text{Number of tosses}}
\end{aligned}$$

Updating Beliefs and RTTM

We can do a little math...

$$\text{Best Guess after } N \text{ tosses} = \frac{H_0 + \#\text{ Heads}}{H_0 + T_0 + N}$$

$$\text{Best Guess after } N \text{ tosses} = \frac{N}{H_0 + T_0 + N} \cdot \frac{\#\text{ Heads}}{N} + \frac{H_0 + T_0}{H_0 + T_0 + N} \cdot \frac{H_0}{H_0 + T_0}$$

$$\text{Best Guess after } N \text{ tosses} = w_N \cdot \text{Proportion of Heads observed} + (1 - w_N) \cdot \text{Initial Best Guess}$$

$$w_N = \frac{N}{H_0 + T_0 + N}, \quad 0 \leq w_N \leq 1$$

Updating Beliefs and RTTM

Recall from earlier:

$$\text{Regressed Estimate} = w \cdot \text{Observation} + (1 - w) \cdot \text{Mean}, \quad 0 \leq w \leq 1$$

Updated belief:

$$\text{Best Guess after } N \text{ tosses} = w_N \cdot \text{Proportion of Heads observed} + (1 - w_N) \cdot \text{Initial Best Guess}$$

$$w_N = \frac{N}{H_0 + T_0 + N}, \quad 0 \leq w_N \leq 1$$

➔ Our best guess can be subjective/opinion (this is *Bayesian Inference*)

➔ Or it can be empirical like the average above (this is *Empirical Bayes*)

Projection Systems

Projection Systems

Many different projection systems: PECOTA, ZiPS, Steamer, Marcel, etc

Lots of similar ideas:

- ➔ Use historical performance
- ➔ Look for similar players to help with forecasting
- ➔ Aging
- ➔ *Regression to the mean*

Marcel Demo

Marcel the Monkey

Win/Loss Records

https://fivethirtyeight.com/features/how-to-predict-mlb-records-from-early-results/

A good primer on Regression to the Mean

One thing to note: which “mean”?

- ➔ Should you regress a team’s W/L record towards .500?
- ➔ Or towards the pre-season projection?
- ➔ (Probably to the projection if it’s good)