

Overall Divergence and Standard Deviation Analysis

Exploratory Research Methodology

1. Overall Score: Player Skill Profile

Overall Score measures a player's statistical skill profile using batting average, on-base percentage, slugging percentage, strikeout rate, and walk rate.

2. Expected Overall

Expected Overall represents the Overall Score associated with players who have the most similar statistical profiles in the current-season population.

Expected Overall = Mean Overall of 10 Nearest Comparable Profiles

Each player's five component scores are standardized, and the 10 closest complete profiles are identified using Euclidean distance.

3. Overall Divergence

Overall Divergence measures how far a player's actual Overall Score is above or below the Overall Score expected from his closest statistical comparables.

Overall Divergence = Overall Score – Expected Overall

Positive divergence means Overall is above expectation; negative divergence means Overall is below expectation; zero means actual Overall equals Expected Overall.

4. Standard Deviation

Standard deviation is calculated from the distribution of Overall Divergence values. Overall Divergence is first calculated for every player, followed by the mean Overall Divergence and standard deviation across the sample.

A player's position in standard-deviation units can then be expressed as:

$$z = (\text{Overall Divergence} - \text{Mean Overall Divergence}) / \text{SD of Overall Divergence}$$

5. Standard Deviation (σ) Bands

Players are grouped according to how far their Overall Divergence falls from the sample mean:

- $< -2\sigma$
- -2σ to -1σ
- -1σ to -0.5σ
- -0.5σ to $+0.5\sigma$
- $+0.5\sigma$ to $+1\sigma$
- $+1\sigma$ to $+2\sigma$
- $> +2\sigma$

6. Research Objective

The purpose of this analysis is to test how the Overall model behaves at different levels of divergence and whether player outcomes differ across the various standard-deviation bands.

The primary outcomes being examined are:

- Average change in Overall Score
- Identity Change rate
- Number of players represented in each standard-deviation band

Research Question

Does the behavior of the Overall model differ depending on how far a player's Overall Score deviates from the Overall Score expected from his closest statistical comparables?

Interpretation

This is an exploratory analysis. The purpose is not to claim that divergence itself predicts an outcome, but to determine whether the model exhibits different behavior at different levels of divergence.

Future analysis can extend this framework across repeated weekly observations to determine whether Overall Divergence and player-value changes show persistent or cyclical patterns over time.