Jacob Norman
ISE501
Dr. King
2024 March 16

<center>Project Proposal</center>

Problem Description

Most people are familiar with one professional sport or another, such as football, hockey, or baseball. Fans routinely choose to engage with these sports through something called fantasy leagues. These are essentially groups of individuals who compete against each other with their own teams. In a fantasy baseball league, these teams are made up of real-life professional players from Major League Baseball (MLB) that the fantasy team owners select. Even within a single type of fantasy sport, there are many different permutations of league rules; however, the performance of a team's players in real-life games is always tracked and then converted into some sort of point system. The goal is to score more points than the other league teams and win the fantasy league at the end of the season.

Of course, there needs to be an equitable way to distribute the player pool, otherwise most teams would end up composed of the same, top performing players and the league would be quite boring. There are two primary means of selecting players: auction drafts and snake drafts. This model will focus on snake drafts, which involve assigning each team a number between 1 and the number of teams in the league. There are then a certain number of rounds, in this case 25, in which players select players in order. The "snake" part of the draft refers to the fact that the team with the last pick in round j has the first pick in round $j + 1$. As an example, in a 12-team league the team with pick 12 would select 12th in the first round, 13th in the second round, 36th in the third round, and so on.

The goal of my MIP model is to select the optimal fantasy team given a set of constraints, such as roster size, minimum position requirements, and average draft position (ADP). The optimal fantasy team will seek to have at least a minimum number in 10 statistical categories using a goal programming approach. These categories include home runs, stolen bases, strikeouts, and wins.

Model Formulation

Sets
i = Player name (~500 players)
j = Round in draft [1-25]
k = Statistical category [HR, R, RBI, SB, OBP, W, SO, SOLD, ERA, WHIP]
p = Position [C, 1B, 2B, 3B, SS, INF, OF, UT, P]

Parameters
T_k = [275, 1000, 1000, 200, 0.35, 100, 1200, 80, 3.7, 1.2]; Target team total for category k
W_k = [1, 1, 1, 1, 1, 1, 1, 1, -1, -1]; Weight for objective k
$Proj_{ik}$ = CSV file; Projected end season total for category k for player i
ADP_i = Average draft position for player i in similar 12-team fantasy drafts
$MinPos_p$ = [1, 1, 1, 1, 1, 5, 4, 11, 7]; Minimum number of players eligible to play position p
Pos_{ip} = CSV file; binary indicator if player i is eligible to player position p
DP = 6; Starting draft position in round 1
$Teams$ = 12; Number of teams in league

Decision Variables

$$x_{ij} = \begin{cases} 1, & \text{if player } i \text{ is selected in the } j\text{th round of the draft} \\ 0, & \text{if player } i \text{ is not selected in the } j\text{th round of the draft} \end{cases}, \forall i, \forall j$$

obj_k^Δ = Normalized deviation of total category k from target k, $\forall k$

Objective Function

$$\max z = \sum_{k=1}^{10} W_k obj_k^\Delta$$

Constraints

Obj)
$$\sum_{i=1}^{500}\sum_{j=1}^{25} \frac{Proj_{ik}x_{ij} - T_k}{T_k} = obj_k^\Delta, \qquad k = 1 - 10$$

RoundMax)
$$\sum_{i=1}^{500} x_{ij} = 1, \qquad j = 1 - 25$$

PlayerMax)
$$\sum_{j=1}^{25} x_{ij} \leq 1, \qquad i = 1 - 500$$

ADPMax)
$$\sum_{i=1}^{500} ADP_i x_{ij} \geq Teams(j-1) + DP, \qquad j = 1 - 25$$

PositionMin)
$$\sum_{i=1}^{500}\sum_{j=1}^{25} Pos_{ip}x_{ij} \geq MinPos_{ip}, \qquad p = 1 - 9$$

$$x_{ij} \text{ bin}, \qquad i = 1 - 500, j = 1 - 25$$
$$obj_k^\Delta \geq 0, \qquad k = HR, R, RBI, SB, OBP, W, SOLD, SO$$
$$obj_k^\Delta \leq 0, \qquad k = ERA, WHIP$$

Numerical Experimentation

For numerical experimentation, I would like to understand how changing certain aspects of the model alter the final roster. Some of these changes include:

- The projections the player statistics ($Proj_{ik}$) are based on between the four main systems;
- The starting draft position in round 1(DP) between several spots;
- Manually selecting certain players in the first few rounds via new constraints; and
- Adjusting logic for when players can be selected based on a confidence interval instead of ADP only.

Other Considerations

This model is based off the league settings for my current fantasy baseball league, not a research paper. Targets were set based on the top three league leaders last season. I know we spoke about the size of the model, but I was able to convert it from a GAMS model to a JuMP model using the Julia programming language.