NHL realignment effects on conference gap

Kilbourne Charrier, Michael Bauers, Clinton, Sam Downs

Motivation

Introduction

Methodology

Basic Idea

The basic idea behind the model is that all of the teams have a "skill level" that they play at or near during any given game. Each game is competitive because we aren't sure how well either team will play, and even a team with a higher skill level can be beaten by another team for any number of reasons. We keep track of this skill level and draw other metrics based off of it, then use the metrics to compute a team's score for the season. This score is used to calculate the team's seed in their respective conference and division.

Creating the Model

To start out, we first need to make sure that we calculate the game-to-game performance levels given by τ and the standard deviation of the team's talent levels given by σ. These are related to the following ratio, where p is the probability that the worst team in the league beats the best team in the league.

$-\frac{2.8854}{\Phi^{-1}(p)} = \frac{\tau}{\sigma}$

Based on the data from all of the seasons with 30 teams, we can calculate that the team with the worst record in the league beat the team with the best team about a quarter of the time. We choose to set p equal to 0.25 for this reason, making the ratio $\tau/\sigma = 4.28$. This means that there are four times the variability in the team's performances in game than there is in the talent levels of the teams. The actual values of τ and σ don't matter, just the ratio, but for simplicity we will set τ equal to 4.28 and σ equal to 1.

We then want to simulate a skill level metric for each of the thirty teams. We create the metric from the following distribution, with team i being equal to μ_i.

$\mu_i \sim N(0, \sigma^2)$

We draw the values from a normal distribution because in sports the overall performace of the teams is generally normally distributed. That is, there are a few really good teams, a few really bad teams, and a lot of teams in between.

We hold the team skill levels constant throughout the season, this make intuitive sense, because there are very seldom large changes to the team's rosters or strategies throughout a season, but after each season teams often make at least some change.

For each game that is played in the season we draw a performance metric for both competing teams. This metric is drawn from a normal distribution with the mean being equal to the team's skill level and the standard deviation being equal to τ^2. This is based on the idea that a team with any skill level will play somewhere around their skill level, rarely playing much worse and rarely playing much better. This metric is calculated as such, with i indicating the home team and j indicating the away team.

$\gamma_i \sim N(\mu_i, \tau^2)$

$\gamma_j \sim N(\mu_j, \tau^2)$

Results