

Big Data: Homework 3

Will Clark & Matthew DeLio
41201-01

April 17, 2015

1 Player Contribution Regression

The model for player contribution is:

$$\log \left[\frac{Pr(y = 1)}{1 - Pr(y = 1)} \right] = \beta_0 + \alpha_{\text{team,season}} + \alpha_{\text{config}} + \sum_{\text{homeplyr}} \beta_j + \sum_{\text{awayplyr}} \beta_j$$

For a given player j , the model estimates the odds multiplier that a goal scored while player j is on the ice was scored by the his team. It includes the following two control factors:

- **team,season:** This should control for high- or low-offense years or certain arenas that provide a special home-ice advantage; and
- **config:** This should control for disproportionate playing time in power plays or end-of-game situations where the goalie has been pulled.

To use one example, the coefficient on Alex Ovechkin is 0.30. This means that a goal scored while Ovechkin is on the ice is $\exp(0.30) = 1.35$ times as likely to be scored by his team, the Washington Capitols, than by their opponents.

We can sort the array of player coefficients to determine the 10 most and least valuable players in the data set. We show the results in Table 1 and Table 2. This evaluation metric accords with our intuition about hockey. The list of 10 best players includes some of the sport's all-time greats, which tells us the model is doing a reasonably good job of quantifying performance.

By this performance metric, the best and worst players are both outliers. We can see in Figure ?? that most players are centered around a 0 rating, and only a small handful of players are significantly better and significantly worse than the mean.

Player	Rank	β_j	$\exp(\beta_j)$
Peter Forsberg	1	0.7548	2.1272
Tyler Toffoli	2	0.6293	1.8762
Ondrej Palat	3	0.6284	1.8746
Zigmund Palffy	4	0.4427	1.5569
Sidney Crosby	5	0.4131	1.5115
Joe Thornton	6	0.3838	1.4678
Pavel Datsyuk	7	0.3762	1.4567
Logan Couture	8	0.3682	1.4451
Eric Fehr	9	0.3677	1.4444
Martin Gelinas	10	0.3578	1.4301

Table 1: Top 10 NHL Players (2002-2014)

Player	Rank	β_j	$\exp(\beta_j)$
Ryan Hollweg	2430	-0.2989	0.7417
Raitis Ivanans	2431	-0.3129	0.7313
Darroll Powe	2432	-0.3340	0.7161
Chris Dingman	2433	-0.3342	0.7159
Mathieu Biron	2434	-0.3512	0.7038
Thomas Pock	2435	-0.3844	0.6809
Niclas Havelid	2436	-0.3855	0.6801
P.J. Axelsson	2437	-0.4284	0.6516
John McCarthy	2438	-0.5652	0.5683
Tim Taylor	2439	-0.8643	0.4213

Table 2: Bottom 10 NHL Players (2002-2014)