The Quantitative Shot Quality Project

xGoals - Glossary and Explanation of Terms

Jamie McCorriston

master_jame_12@hotmail.com

Twitter: @masterjame12

Connor Reed

connor.reed.92@gmail.com

Last Updated: January 21, 2016

Table of Contents

A New Terms Relevant to Our Study

A.1 xGoal

Rather than discretely summing a goal, shot attempt (Fenwick), or shot attempt+blocked shot attempt (Corsi), we have created the ***xGoal***. By graphing shooting percentage (y-axis) against shot attempt distance (x-axis), we found a strong relationship between the two (as one might expect). From this, we smoothed the function, which resulted in a formula taking distance as an input, and outputting an expected shooting percentage for that distance, or, an xGoal.

You can think of it like this: A shot from 5 feet away from the goal may be expected to result in a goal 20% of the time. So when a player takes a shot from this distance, he is credited with 0.2 xGoals. However, a shot taken from 35 feet away from the goal may result in a goal 5.0% of the time. So when a player takes a shot from this distance, he is credited with 0.050 xGoals. By taking each shot, plugging it into our function, and summing the results for each player, we can calculate the players expected goals for a time period based on the shots they took. Going back to our example, say a player took 1 shot from 5 feet away, and 5 shots from 20 feet away. Under our model, he will have amassed $((1 \times 0.2) + (5 \times 0.050)) = 0.45$ xGoals regardless of how many real goals he scored over this stretch. This is the general outline of how we have calculated xGoals.

A.1.1 Disclaimer

Our model is only dictated towards shots taken in a 5v5 or 5v4 situation. Therefore, all shots and goals scored at 5v3, 4v5, 4v4, 4v3, 3v5, 3v4, 3v3, or penalty shot situation have been omitted, as we do not have adequate data to determine an average shooting percentage in terms of shot distance for these situations.

Additionally, it should be noted that this model ***does not*** take into account the ***location*** of a shot, it only takes into account the ***distance*** (so a shot taken from the goal line 5 feet away from the goal counts the same as a shot taken from directly in front of the goal 5 feet away). There is also no passing data that has been used, so in terms of shot quality, the only variable we are measuring is distance, not location, or quality of play up until the time of the shot.

A.1.2 Goals Above Expected (GAE)

Clearly, some players are better at shooting than others (see Stamkos, Steven or Kovalchuk, Ilya). For this, we would like to see how many more real goals they amassed over what the model would suggest. By

This is done by simply subtracting a players expected goals (xGoals) value from their actual goal total. For example, in the year 2013/14, Alex Ovechkin led the league in 5v5 and 5v4 goals scored with 43. Over the same span, Ovechkin's 501 shot attempts resulted in what we would have expected to produced 28.699 goals. Therefore, his Goals Above Expected (GAE) value stands at $(43 - 28.699) = +14.301$.

This metric is positively bounded by the amount of goals a player has scored, and is negatively bounded by the amount of goals we would expected a player to have scored (xGoals). It can be seen that the magnitude of this metric is largely determined by time on ice (and thus shot attempts).

A.1.3 Goals Per Expected (GPE)

To normalize GAE for shots taken, we will instead divide a players actual goals by their expected goals.

Again, using Alex Ovechkin's 2013/14 season as an example. We take his 43 5v5 and 5v4 goals, and his 28.699 expected goals, to get:

$$\frac{\text{43 actual goals}}{\text{28.699 expected goals}} = \text{1.498 Goals per Expected Goal (GPE)}$$

Therefore, we see that in 2013/14, Alex Ovechkin potted nearly 1.5 goals for every goal we would expect him to score based on the location of his shot attempts. Is this a skill, or is this luck? Well, from 2007-present (early January, 2015), Alex Ovechkin has a GPE value of 1.355. That is, he has scored 1.355 times as many goals as we would expect over this time period, so while his GPE of 1.498 in 2013/14 may have been outside of his career average, it is not so far away for us to expect it to be too much of an outlier value.

As a rate statistic, this metric contains a lower bound of 0, as it would be impossible for a player to record negative goals, or negative expected goals.

A.2 xDifference

A common statistic in the up and coming hockey analytics community has been Fenwick. Named after Calgary Flame fan and blogger Matt Fenwick, this is simply a measure of unblocked shot attempts. By weighting each of these attempts in terms of expected goals, we gather a metric for weighted Fenwick, which we have called the expected goal differential percentage, or xDiff.

For example, say that while a player is on the ice, his team gathers 1000 unblocked shot attempts for, 94 goals for, and 95.5 xGoals for, but they allow 900 unblocked shot attempts, 111 goals against, and 104.5 xGoals against. From these values, we can compute both goals for% (GF%) and Fenwick for% (FF%), and using the xGoal values we listed, it is also possible to compute expected goal differential% (xDiff%).

- Goals For (GF%)

$$\frac{\text{94 goals for}}{\text{94 goals for} + \text{111 goals against}} = \frac{\text{94 goals for}}{\text{205 total goals scored}} = 45.85\ \text{GF\%}$$

- Fenwick For (FF%)

$$\frac{\text{1000 shot attempts for}}{\text{1000 shot attempts for} + \text{900 shot attempts against}} = \frac{\text{1000 shot attempts for}}{\text{1900 total shot attempts}} = 52.63\ \text{GF\%}$$

- Expected Goal Differential Percentage (xDiff%)

$$\frac{\text{95.5 xGoals for}}{\text{95.5 xGoals for} + \text{104.5 xGoals against}} = \frac{\text{95.5 xGoals for}}{\text{200 total xGoals}} = 47.75\ \text{xDiff\%}$$

A.2.1 Adjusted xDifference

By adjusting for where players start their non-open play shifts (i.e., where their shifts that begin with a faceoff occur, either the defensive zone, neutral zone, or offensive zone), we can slightly adjust the expected goal differential value. Rarely does this make a dramatic impact, but in the rare case of defensive zone specialists (see Malhotra, Manny and Gordon, Boyd) or offensive zone specialists (see the Sedin twins in Vancouver or the Patricks [Kane and Sharp] in Chicago) it can move the xDiff needle a little bit.

A.3 Team xDifference

So, how can we use our model to help gauge team-wide performance? Well, looking past the W-L column, one of the easiest ways to gauge the performance of a team as a whole is to compare the amount of goals they've scored to the amount of goals they've allowed. We can do similarly with the amount of goals we ***expect*** a team to score and allow. By dividing the amount of xGoals for, by the total amount of xGoals, we get a percentage of the goals we would expect a team to score in a given time frame. This is what we call the expected difference, or **xDiff** of a team.

Again, let's do an example. Looking at our data in 2013/14, the leader in team xDiff was the San Jose Sharks. We credited them with 237.276 xGoals, while allowing only 173.664 xGoals. Therefore, we find that they will be expected to have:

$$\frac{237.276 \text{ xGoals for}}{237.276 \text{ xGoals for} + 173.664 \text{ xGoals against}} = \frac{237.276 \text{ xGoals for}}{410.940 \text{ total xGoals}} = 57.73 \text{ xDiff } \%$$

It should also be noted that if we subtract (power play xGoals for) from the numerator, and (power play xGoals for +penalty kill xGoals allowed) from the denominator, we can calculate even strength xDiff, or **evxDiff**

$$\frac{237.276 \text{ xGoals for} - 61.889 \text{ 5v4 PP xGoals for}}{410.940 \text{ Total xGoals} - 61.889 \text{ 5v4 PP xGoals for} - 34.883 \text{ 4v5 PK xGoals against}} = 55.83 \text{ evxDiff } \%$$

A.4 Adj. Save% (for Goalies)

Since we have a measure for the quality of shots players and teams are amassing, we may as well use it to measure the types of shots different goalies will have to face. As a whole, we would expect all goalies to face a relatively similar quality of shot over a large sample, as they will all face the same opponents (with the exception of a goalie will never face his own team). However, there are some drivers that would alter the quality of shots that a goalie would face (defensive system of their own team, playing in a particularly strong division, etc.), so this is not exactly a trivial pursuit.

As the most easily accessible and well recognized measures for goalie performance, save percentage is relatively easy to compute:

$$\text{Sv\%} = \frac{\text{Shots Faced - Goals Allowed}}{\text{Shots Faced}}$$

By replacing the "Goals Allowed" portion of the numerator with xGoals against, we find expected Sv%.

$$\text{Expected Sv\%} = \frac{\text{Shots Faced - xGoals Againt}}{\text{Shots Faced}}$$

For this example, let's look at 2013/14 Vezina Trophy winner Tuukka Rask. In situations we analyzed, Rask faced 1548 shots against, allowed 103 goals, while being expected to allow 117.640 goals. We can compute his Sv%, and expected save percentages as follows:

$$\text{Sv } \% = \frac{1548 - 103}{1548} = 93.34\% \qquad \text{Exp. Sv\%} = \frac{1548 - 117.640}{1548} = 92.40\%$$

Next, we look at the league wide expected save percentage against, and we see that Rask's Exp. Sv% is higher than league average (essentially, the shots he faced were less dangerous than most goalies). Therefore, we must make an adjustment to his actual Sv%. In this case, the adjustment is a -0.5% adjustment, and Rask's Adjusted Save percentage in 2013/14 is 92.8%, good for 9th in the league amongst goalies facing over 1000 shots.

Essentially, the adjusted Sv% metric gives goalies a small boost up or down depending on the quality of shots they've faced. If you faced more dangerous shots than average, you get a small bump up, and if you face less dangerous shots like 2013/14 Tuukka Rask, you get a small bump down.

B A Non-Comprehensive List of Important Websites to Visit

B.1 Glossaries

PLEASE NOTE: There is a lot of crossover between these sites.

Sites are listed alphabetically by section.

B.1.1 The original BehindTheNet (BTN) glossary

http://www.behindthenet.ca/stats_faq.php

B.1.2 A more in-depth, less comprehensive glossary from BroadStreetHockey (BSH)

http://www.broadstreethockey.com/pages/bsh-advanced-stats-glossary

B.1.3 The wonderful glossary created by David Johnson of Puckalytics (and Hockey-Analysis)

http://www.puckalytics.com/glossary.html

B.1.4 Eric Tulsky, formerly of BroadStreetHockey, wrote a similar glossary to BSH for the SBNation 2013 season preview, which gives good summaries of the basic statistics being used

http://www.sbnation.com/nhl/2013/9/30/4784288/nhl-season-preview-2013-14#38-

B.2 Advanced Stat/Hockey Analytics/#Fancystat Primers

PLEASE NOTE: Again, there is a lot of crossover between these sites.

Sites are listed alphabetically by section.

B.2.1 Hawerchuk (also the manager of BTN) wrote arguably the best primer on the web for ArcticIceHockey

http://www.arcticicehockey.com/2009/10/9/1078607/frequently-asked-questions-about

B.2.2 A complete listing of his "Understanding Advanced Stats" series are here, with 3 additional posts

http://www.arcticicehockey.com/advanced-stats/archives

B.2.3 The crew at HabsEyesOnThePrize wrote a similar primer more recently. This is especially good if your mother tongue is French, as they have kindly translated some of the articles into French

http://www.habseyesontheprize.com/2013/7/16/4529774/fancy-stat-summer-school

B.3 Eric Tulsky's Reference Libraries

PLEASE NOTE: These libraries are a little outdated (i.e. last updated in 2012), but are still very useful and relevant. Unfortunately, only 4 of 11 topics were successfully libraried before by the project went silent and eventually the hockey-centric portion of Tulsky's mind was hired away by an NHL organization http://nhlnumbers.com/2012/11/1/stats-article-reference-library

B.3.1 While you're at it, you may as well read up on all of Eric's work at NHLnumbers and at his Outnumbered blog on the SBNation

http://nhlnumbers.com/2012/11/1/stats-article-reference-library

http://www.sbnation.com/outnumbered