Sid the Kid's value, based on the Pythagorean theorem

November 19, 2011 - 4:38am By JASON BROWN

As of this moment, Sidney Crosby, the Pittsburgh Penguins' star player, is still out of action after a pair of concussions last January.

I still wince when I see video of the Washington Capitals' David Steckel's blindsided shoulder hit to Sidney's head.

The loss of Crosby has not only affected the Penguins, but hockey itself. I think while there is plenty of blame to go around, much of it lands squarely on the shoulders of the NHL administration that has tolerated such dangerous hits to the head.

Only now does it seem that the league is becoming tougher on offenders, but for Crosby, it may be too little, too late.

Is there a way to quantitatively evaluate just how costly the loss of Crosby is to the Penguins? Bill James, a statistician who has applied mathematics to baseball (a field of study called sabermetrics), put forward a formula for how to estimate the percentage of games won by a baseball team in a season from the number of runs scored (S) and the number of runs allowed (A):

$$\text{Games won} = 82 \times \frac{S^2}{S^2 + A^2}$$

If you think back to school, you might recognize the denominator of the fraction as looking like the Pythagorean formula for calculating the length of the square hypotenuse from the sum of the squares of the other two sides.

Because of this similarity, the formula is often called the Pythagorean winning percentage. If you look hard at the formula, you'll see that increasing the runs scored (or decreasing the runs allowed) increases the winning percentage, as it should.

On the surface it sounds like it would be impossible to predict the percentage of games won from the runs scored and runs allowed, as in any given game, you could score many more runs than you need to win, and yet get only one win, and in another game you could lose badly and still it's only one loss.

Yet we are talking not about a single game, but the sum total of all the games in a long season, and this is where statistics come to the fore.

The formula actually works quite well for baseball teams, giving results that are almost always within five per cent of the true value.

The formula can be carried over to many other sports, such as basketball, football, and indeed hockey. So here is a Pythagorean formula for predicting the number of games won by a hockey team during an 82 game season, based on the number of goals scored (S) and goals allowed (A):

$$\text{Winning percentage} = \frac{S^2}{S^2 + A^2}$$

(I ignored tie games, but the formula seems to work well enough as is.)

For example, in the 2006-07 season, the Penguins scored 267 goals and allowed 240, with 47 games won.

The Pythagorean formula would predict, just based on the goals scored and allowed, that they would have ended up winning 47 games, so we are spot on with the prediction.

Now, how did Sidney's 120 points contribute to the Penguin's success?

Well, without those 120 goals, the Pythagorean formula suggests the team would only have won about 22 games, so having Sidney on the ice added about 25 wins to the season!

In the 2006-07, 2008-09, 2009-2010 seasons (I didn't include Sidney's rookie year or the ones when he played significantly less than 82 games), the formula predicts that in each of these seasons, just based on his offensive output, Sidney added at least 20 wins per season to the Penguins.

In hockey's valuable commodity of wins, Sid the Kid is a top broker — his scoring is the difference between the Penguins being at the ceiling or the basement of the league.

Even the math shows that professional hockey has to get its head together.

Jason I. Brown is a professor of mathematics at Dalhousie University in Halifax.

About the Author »

By JASON BROWN