

## Introduction

前段时间去达特茅斯看费德勒的时候，听到他演讲里说，自己比赛的胜率将近80%，但是每一分的得分概率只有54%。我当时第一反应是有点反直觉，但后来仔细想想，发现还是数学学的不好，自己的直觉调教的还不够好。结合网球的比赛赛制，这个得分概率达到80%的胜率，其实是完全不意外的，在伯努利假设下，这个比赛胜率是偏低的。

太长不看版：假设每一分都是伯努利试验，以0.54概率赢球，那么在三局两胜下，费德勒赢下比赛的胜率为0.859，五局三胜则会提高到0.910。背后的数学直觉是，比赛并不是一分定胜负，即使赢球的概率只是略高于50%，只要比的足够多，最后赢下比赛的概率就会逼近1。

## Assumption

首先再描述一下网球比赛的规则，网球比赛有point（分）、game（局）、set（盘）、match（赛）四个阶段，费德勒的match胜率有80%左右，但是point胜率只有54%。

在一局中，先获得4分（并且领先2分）的一方获得胜利；在一盘中，先赢6局（并且领先2局）的一方获得胜利，如果打到6:6，则会进行抢七，先赢7分（并领先2分）者获得该盘胜利；在一场比赛中，根据赛制，可能会有5局3胜或3局2胜。

事实上抢七规则比较复杂，不同赛事的规定会有细微的差别，例如在一般的大满贯赛事中，只有前四盘进行抢七，第五盘会一直打到一方领先两局；但是us open允许第五局进行抢七。为统一模型，我们假设所有盘均进行抢七。

同时，发球局和接发球局的得分胜率肯定是有区别，但我们这里不做区分，统一假设一分为一个伯努利分布，费德勒会以 $p$ 的概率拿下这一分。

## 局胜率

首先计算局胜率  $p_g$ 。分成不加赛和加赛两部分计算。

不加赛的部分，可能的结果是4:0，4:1，4:2，同时必须赢下最后一分，于是可能的组合数为  $C_3^0, C_4^1, C_5^2$ ，对应的概率为  $C_{k-1}^{k-n} p^n (1-p)^{k-n}$ ，这里  $n = 4, k = 4, 5, 6$ 。

另一种可能是打到了加时，双方一直在比拼谁能领先两分。打到加时的前提条件为3:3。进入加时之后，我们可以将这个过程看作随机游走，以 $p$ 的概率加一， $1-p$ 的概率减一，问题就转化为，从0开始游走，在达到-2之前先达到2的概率是多少。对此可以用动态规划的思路来求解。假设从 $k$ 出发，在达到-2之前先达到2的概率为 $a_k$ ，那么我们有

$$\begin{cases} a_2 = 1 \\ a_1 = pa_2 + (1-p)a_0 \\ a_0 = pa_1 + (1-p)a_{-1} \\ a_{-1} = pa_0 + (1-p)a_{-2} \\ a_{-2} = 0 \end{cases}$$

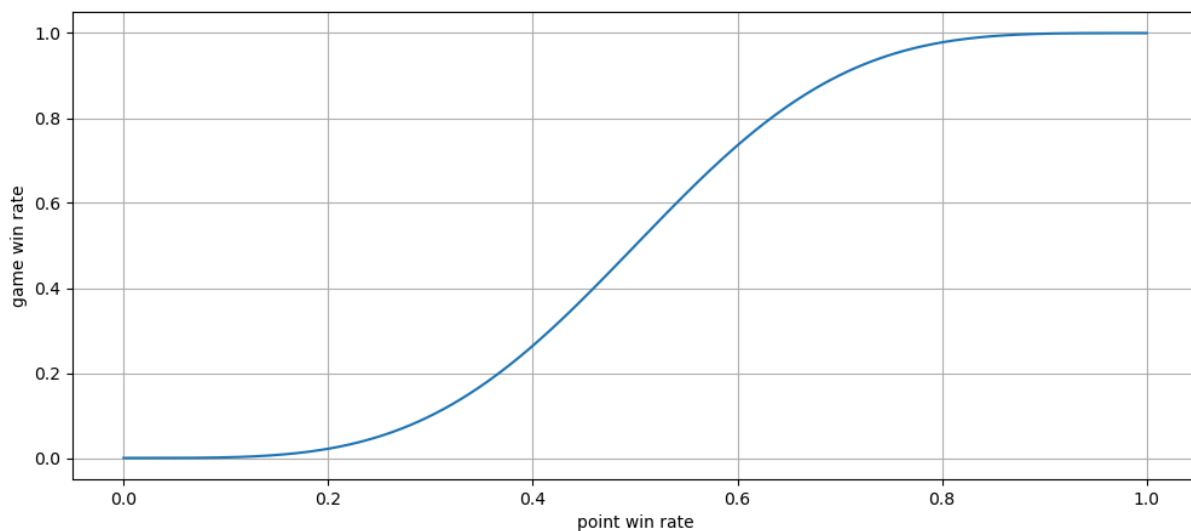
求解该线性方程，得到

$$a_0 = \frac{p^2}{1 - 2p(1 - p)}$$

综上，我们可以给出局胜率对分胜率的公式

$$p_g = \sum_{k=n}^{2n-2} C_{k-1}^{k-n} p^n (1-p)^{k-n} + C_{2n-2}^{n-1} p^{n-1} (1-p)^{n-1} a_0, n = 4$$

对应图像大致如下



分胜率为0.54时，对应的局胜率为0.599。同时从图中可以看到，当分胜率为0.8时，局胜率就已经逼近1了。

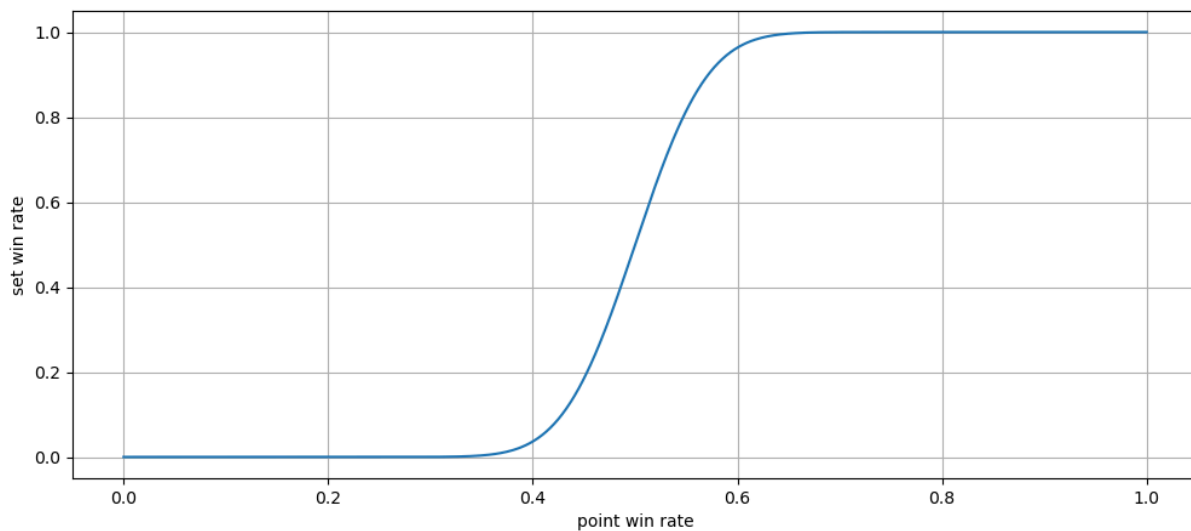
## 盘胜率

然后基于局胜率 $p_g$ 计算盘胜率 $p_s$ 。注意到，将局胜率公式中的 $n$ 从4替换为7，就是抢七的胜率 $p_7$ 。

直接拿下比赛，可能的情况有6:0, 6:1, 6:2, 6:3, 6:4，这与 $p_g$ 公式中的第一部分相同。或者进入5:5，可能的情况有7:5拿下这盘，或者进入6:6打抢七赢下比赛。因此盘胜率的计算公式为

$$p_s = \sum_{k=n}^{2n-2} C_{k-1}^{k-n} p_g^n (1 - p_g)^{k-n} + C_{2n-2}^{n-1} p_g^{n-1} (1 - p_g)^{n-1} (p_g^2 + 2p_g(1 - p_g)p_7),$$

其中 $n = 6$ 。图像绘制如下



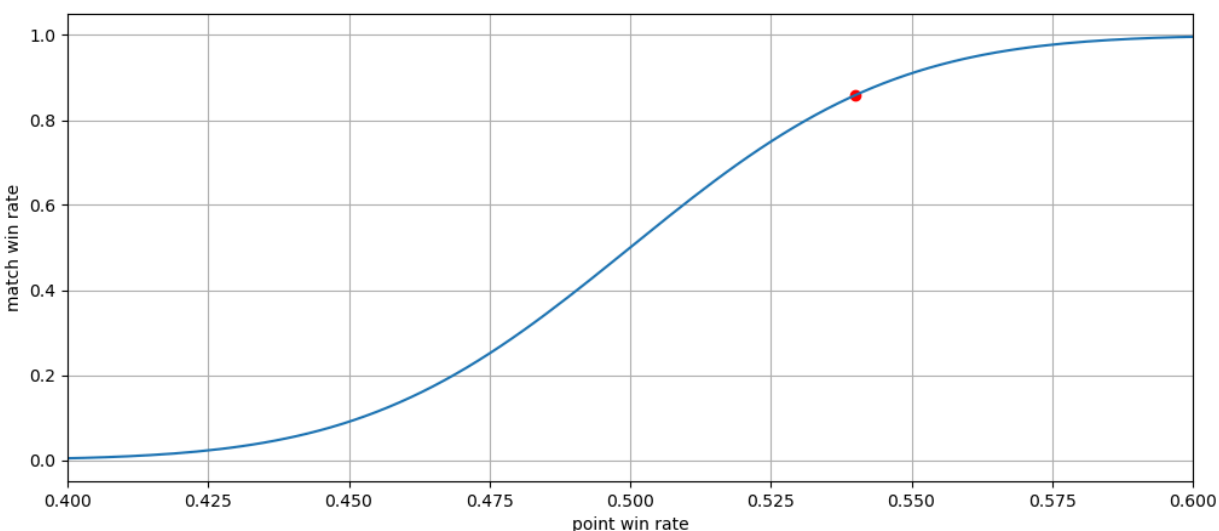
此时，0.54的分胜率对应的盘胜率为0.763。从图中也可以看到，当分胜率达到了0.6时，盘胜率已经接近1。

## 比赛胜率

下面基于盘胜率 $p_s$ ，我们来计算比赛胜率 $p_m$ 。这同样类似计算局胜率的第一部分，仅有细微差别我们直接给出结果

$$p_m = \sum_{k=n}^{2n-1} C_{k-1}^{k-n} p_s^n (1 - p_s)^{k-n}$$

其中 $n = 2$  或  $3$ 。以三局二胜为例，对应图像为



注意这张图中我把横轴调宽了，其中红点是费德勒所在的位置。三局两胜下的比赛胜率为0.859，而五局三胜的话就会提高到0.910。

## Conclusions

网球赛制放大了球员之间的差异。表面上，54%是一个很低的胜率，但其实很高。业余球员有时也有类似的感觉，对手只比我强一点，但是set就是打不赢。

我们也不难想到，把胜率从50%提升到54%，背后需要付出多少的努力，尽管数字上的提升并不算多。