

Почему мы неправильно считаем в спортивном бридже, или «нормальная» игра в бридж.

Немного очевидностей. Как мы считаем match point в бридже? Выделяют два алгоритма подсчета:

- «Игра на среднее» (imp)
- «Игра на Макс»

Разберемся, как устроены эти подсчеты.

Подсчет match point при «Игре на среднее».

Давайте договоримся о следующих обозначениях:

- N – Количество записей в протоколе
- x – Протокол. Набор записей в протоколе.
- x_i – Результат

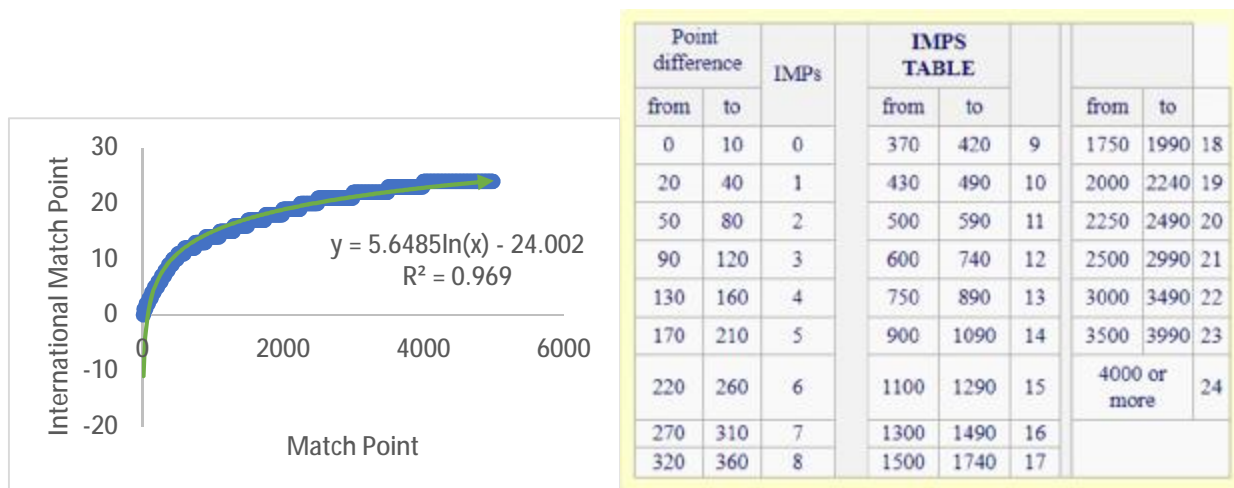
При игре на среднее определяется показатель «цена сдачи», как среднее арифметическое всех записей¹:

$$\mu = \frac{1}{N} \sum x_i$$

Тогда результат пары – это отклонение полученного результата в сдаче от цены сдачи. Такой результат получил название match point:

$$mp_i = x_i - \mu$$

Далее mp переводят по какой-то совершенно дикой формуле в так называемые international match point (imp). Смысл этих преобразований в том, чтобы избавиться от диких чисел, и немного сгладить большие выигрыши. Саму формулу мне найти не удалось, но вот что подсказывает регрессионный анализ таблицы imp'ов.



¹ Не будем сейчас вдаваться в подробности округления расчетов, а также проблемы борьбы с «выбросами» и использование урезанного среднего.

Подсчет match point при «Игре на МАКС».

При «Игре на Макс» используется немного другой алгоритм. Сначала все записи упорядочиваются по возрастанию, и каждой записи присваивается ранг r_i – порядковый номер. В случае равенства записей, все такие записи получают среднеарифметический ранг.

Ценой сдачи принимается среднее арифметическое всех рангов:

$$\bar{r} = \frac{1}{N} \sum r_i = \frac{N+1}{2}$$

Тогда match point в этой игре присваивается по формуле:

$$mp_i = r_i - \bar{r}$$

Заметим, что формула начисления match point в «Игре на Макс» аналогична формуле для начисления match point в «Игре на среднее», только вместо самих значений используется ранг записи. У этой формулы есть и еще один недостаток – она зависит от количества записей в протоколе, и если случается так, что какую-то сдачу надо исключить из расчетов (сдача «испорчена»), то результаты становятся несравнимыми. Для решения этой проблемы применяют корректировку, а результат выражают в процентах:

$$\% = \frac{1}{2} + \frac{mp_i}{N-1}$$

Теперь, в отличие от «Игры на среднее», осталось не просуммировать mp , а усреднить %, полученные парой в каждой сдаче.

«Нормальная игра»

Существование двух разновидностей игры в бридж приводит к разным стратегиям игры. Например:

- При «Игре на среднее» становится важным ставить и реализовывать гейма и шлема, в то время как при «Игре на Макс» становится существенной зона частичной записи.
- При игре на среднее записи 430 и 420 имеют разницу в $0mp$, а при «Игре на Макс» это разные ранги.
- Разные стратегия виста – при «Игре на Макс» не столь важно посадить контракт, сколь не дать лишнюю взятку. Ну и так далее...

Но самый главный недостаток тот, что все сдачи разные, и в каждой сдаче решаются разные проблемы, и стоимость этих проблем разная. Это означает, что сами по себе сдачи между собой не сравнимы, а значит и суммировать полученные mp не корректно. Что делать?

Помните анекдот про то, что средняя температура по больнице составляет 36.6°C. Анекдотичность ситуации возникает из-за того, что если замерять температуру у здоровых людей, то это значение не сильно будет отличаться от средней. А вот у реальных людей в больнице она сильно должна отличаться. Так появляется мера отклонения от середины.

В качестве такого показателя использует понятие дисперсии, или среднеквадратического отклонения²:

$$D = \frac{1}{N} \sum (x_i - \mu)^2$$

² Эту формулу можно упростить: $D = \frac{1}{N} \sum x_i^2 - \mu^2$

Или стандартного отклонения³:

$$\sigma = \sqrt{D}$$

Чем выше отклонение реальных данных от середины, тем выше S . А если все данные одинаковые, то $\sigma = 0$.

Чем сложнее проблема, тем больше будет разброс записей в протоколе. Чем легче проблема – тем меньше будет разброс.

Вот теперь мы можем все сдачи уравнивать с точки зрения разброса, и сказать, что для победы важны все проблемы, и они имеют одинаковый вес. Match Point в такой системе вычисляются по следующей формуле:

$$mp_i = \frac{x_i - \mu}{\sigma}$$

Для этой формулы не важно, как при подсчете в «Игре на среднее», количество записей, но важно обнаружение проблемы в сдаче и ее решение. Все сдачи становятся в статистическом смысле одинаковыми. И никаких дополнительных преобразований не требуется. Ценой сдачи мы будем называть пару (mp). Результат турнира есть простая сумма mp .

Такая процедура хорошо известна в мире статистики и методах извлечения знаний из данных. Она называется «нормализацией» данных в том смысле, что у величины:

$$z_i = \frac{x_i - \mu}{\sigma}$$

среднее значение равно нулю, а мера разброса $\sigma = 1$.

Именно так корректно обрабатывать статистическими методами протокол в спортивном бридже, а такую разновидность игры предлагаю назвать «Нормальной игрой», или «Нормализованной игрой», или «Z-игрой», или «Игрой на разброс (стандартное отклонение)».

В практическом смысле интересно округлять расчеты Z до двух знаков после запятой. Это позволит несколько сгладить игру.

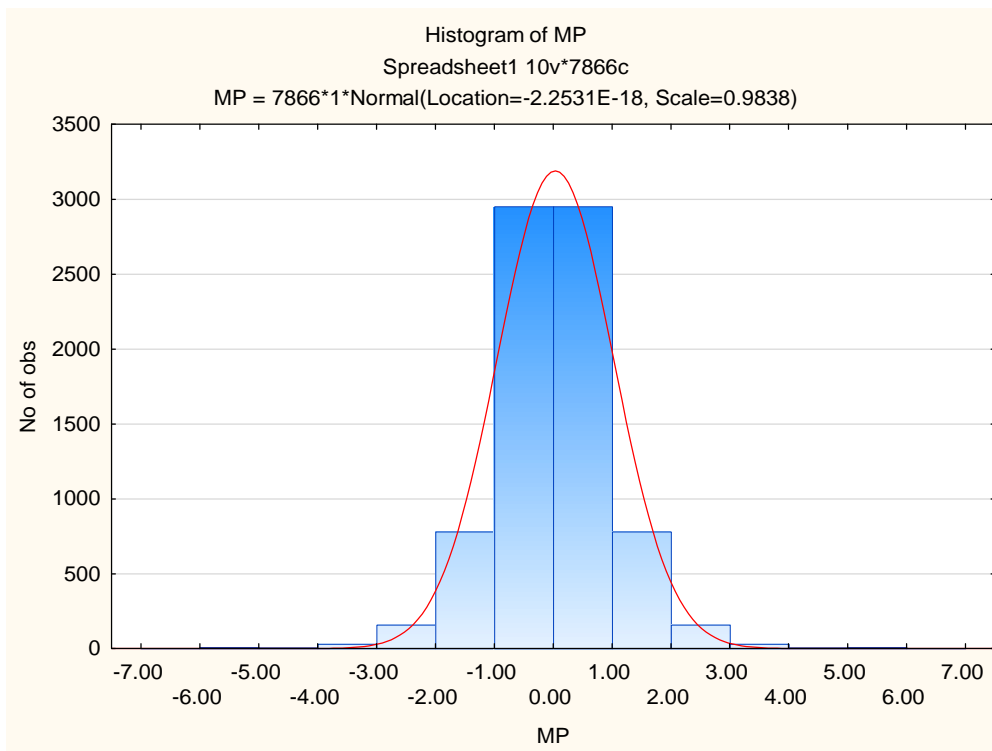
Вероятностное распределение «стандартизированных» match points

В предлагаемой методике расчета в каждой сдаче разыгрывается одинаковое количество mp (в статистическом смысле). Поэтому интересно посмотреть на распределение mp по набору, состоящему из всех сдач.

Обработка реальных данных показывает, что в итоге мы имеем нормальное распределение с параметрами $N(0,1)$. Что можно проинтерпретировать как то, что игрок в каждой сдаче получает mp с нормальным вероятностным распределением.

³ Приведенная выше формула дает так называемое смещение, и поэтому на практике используют формулу

$$\sigma = \sqrt{\frac{N}{N-1}D}$$



Особенности «Нормальной игры» в командном чемпионате

В командном турнире, когда одна команда играет против другой и в протоколе есть только две записи приведенная формула вырождается в простой подсчет: ± 1 , где +1 – это победа, -1 – поражение, 0 – ничья. Судя по наблюдению в ВВО, очень распространенная игра.

Однако если играется команда против команды, но эту же сдачу играют и другие команды, то есть в протоколе сдачи более одной записи, то расчет нормализации может дать более интересные результаты.

Заметим, что в традиционном командном турнире match point вычисляется как разница между результатами пар, находящихся в одной команде, но сидящих на разных линиях NS и EW. А в случае наличия понятия стоимости сдачи μ , эту величину можно добавить и вычесть:

$$mp = x_{NS} - x_{EW} = (x_{NS} - \mu) + (\mu - x_{EW})$$

Если для каждой сдачи мы можем высчитать цену сдачи: (mp_s), то результат команды будет сумма match point для двух пар, сидящих на разных линиях:

$$mp = mp_{NS} + mp_{EW}$$

Таким образом, командный турнир – это разновидность парного турнира, в которой за одну команду играет две пары, одна из которых всегда сидит на NS, а другая на EW, при чем они не играют друг против друга.

Примеры:

Пример №1

N= 12		Игра на среднее		Игра на MAX			Нормализация	
N	Result	mp	imp	Rank	mp	%	Z	N(0,1)
1	450	32	1	12.0	5.5	100.00%	2.00	98%
2	430	12	0	10.0	3.5	81.82%	0.74	77%
3	430	12	0	10.0	3.5	81.82%	0.74	77%
4	430	12	0	10.0	3.5	81.82%	0.74	77%
5	420	2	0	6.5	0.0	50.00%	0.11	54%
6	420	2	0	6.5	0.0	50.00%	0.11	54%
7	420	2	0	6.5	0.0	50.00%	0.11	54%
8	420	2	0	6.5	0.0	50.00%	0.11	54%
9	400	-18	-1	2.5	-4.0	13.64%	-1.16	12%
10	400	-18	-1	2.5	-4.0	13.64%	-1.16	12%
11	400	-18	-1	2.5	-4.0	13.64%	-1.16	12%
12	400	-18	-1	2.5	-4.0	13.64%	-1.16	12%
$\mu=$	418	0	0	6.5	0.0	50.00%	0.00	50%
$s=$	16	16	1	3.5	3.5	31.37%	1.00	30%

Пример №2

N= 12		Игра на среднее		Игра на MAX			Нормализация	
N	Result	mp	imp	Rank	mp	%	Z	N(0,1)
1	430	185	5	12.0	5.5	100.00%	1.44	93%
2	420	175	5	10.5	4.0	86.36%	1.36	91%
3	420	175	5	10.5	4.0	86.36%	1.36	91%
4	400	155	4	9.0	2.5	72.73%	1.21	89%
5	170	-75	-2	6.0	-0.5	45.45%	-0.58	28%
6	170	-75	-2	6.0	-0.5	45.45%	-0.58	28%
7	170	-75	-2	6.0	-0.5	45.45%	-0.58	28%
8	170	-75	-2	6.0	-0.5	45.45%	-0.58	28%
9	170	-75	-2	6.0	-0.5	45.45%	-0.58	28%
10	150	-95	-3	2.5	-4.0	13.64%	-0.74	23%
11	150	-95	-3	2.5	-4.0	13.64%	-0.74	23%
12	120	-125	-3	1.0	-5.5	0.00%	-0.97	17%
$\mu=$	245	0	0	6.5	0.0	50.00%	0.00	47%
$s=$	128	128	4	3.5	3.5	31.49%	1.00	33%

Пример №3

N= 12		Игра на среднее		Игра на MAX			Нормализация	
N	Result	mp	imp	Rank	mp	%	Z	N(0,1)
1	420	252	6	11.5	5.0	95.45%	1.46	93%
2	420	252	6	11.5	5.0	95.45%	1.46	93%
3	400	232	6	10.0	3.5	81.82%	1.35	91%
4	170	2	0	8.0	1.5	63.64%	0.01	50%
5	170	2	0	8.0	1.5	63.64%	0.01	50%
6	170	2	0	8.0	1.5	63.64%	0.01	50%
7	140	-28	-1	5.0	-1.5	36.36%	-0.16	43%
8	140	-28	-1	5.0	-1.5	36.36%	-0.16	43%
9	140	-28	-1	5.0	-1.5	36.36%	-0.16	43%
10	-50	-218	-6	2.0	-4.5	9.09%	-1.27	10%
11	-50	-218	-6	2.0	-4.5	9.09%	-1.27	10%
12	-50	-218	-6	2.0	-4.5	9.09%	-1.27	10%
$\mu=$	168	0	0	6.5	0.0	50.00%	0.00	49%
$s=$	172	172	4	3.5	3.5	32.02%	1.00	30%

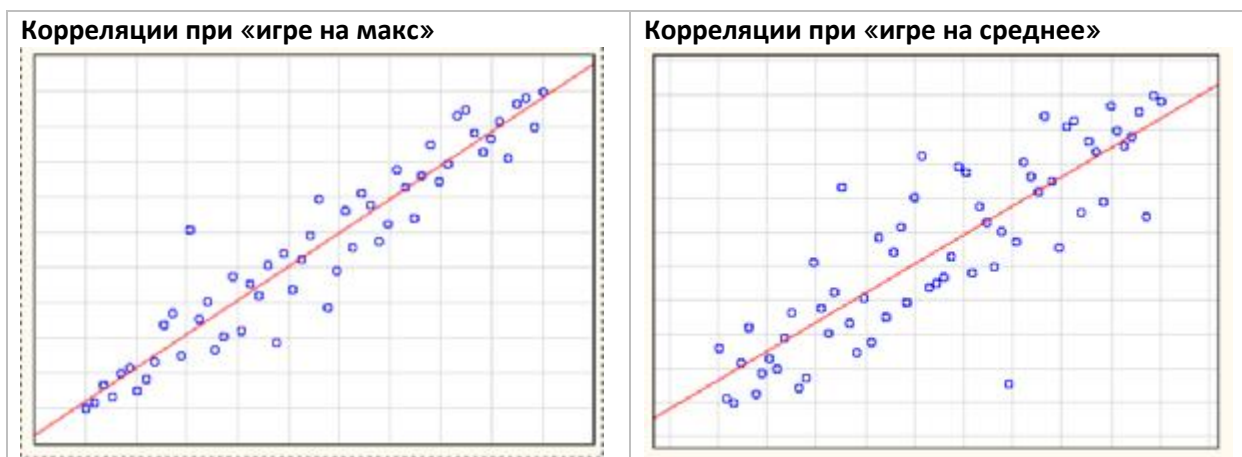
Пример №4

N= 12		Игра на среднее		Игра на MAX			Нормализация	
N	Result	mp	imp	Rank	mp	%	Z	N(0,1)
1	800	348	8	12.0	5.5	100.00%	3.18	100%
2	420	-32	-1	6.0	-0.5	45.45%	-0.29	39%
3	420	-32	-1	6.0	-0.5	45.45%	-0.29	39%
4	420	-32	-1	6.0	-0.5	45.45%	-0.29	39%
5	420	-32	-1	6.0	-0.5	45.45%	-0.29	39%
6	420	-32	-1	6.0	-0.5	45.45%	-0.29	39%
7	420	-32	-1	6.0	-0.5	45.45%	-0.29	39%
8	420	-32	-1	6.0	-0.5	45.45%	-0.29	39%
9	420	-32	-1	6.0	-0.5	45.45%	-0.29	39%
10	420	-32	-1	6.0	-0.5	45.45%	-0.29	39%
11	420	-32	-1	6.0	-0.5	45.45%	-0.29	39%
12	420	-32	-1	6.0	-0.5	45.45%	-0.29	39%
$\mu=$	452	0	0	6.5	0.0	50.00%	0.00	44%
$s=$	110	110	3	1.7	1.7	15.75%	1.00	18%

Корреляции

Шехов П.В, пересчитал турнир «на макс» в стандартизованные match point, а я пересчитал турнир «на среднее». Вот какие корреляции получились:

	Spearman R	Kendal Tau
Турнир «на Макс»	95.24%	82.81%
Турнир «на среднее»	84.73%	67.11%



Таким образом, несмотря на то, что природа расчетов похожа на «игру на среднее», характер игры более похож на «игру на макс».

Судейские проблемы

Предлагаемый подход позволяет легко решать судейские проблемы:

1. Необходимо присудить штраф.

В судейской практике принято присуждать штраф при игре «на среднее» в виду 2 imp, или 60%-40% при игре на среднее. Почему так – сложно сказать – так принято. Поскольку распределение стандартизованных mp нормальное, то 60% составляет 0.253347mp, или, если округлять, 0.25mp.

2. Разное количество записей в протоколах

Поскольку цену игры (ms) мы оцениваем через средние величины, то уменьшение количества записей, как и при игре «на среднее» не оказывает влияние на расчет match point.

3. Разное количество сыгранных сдач

Если какие-то пары сыграли меньшее количество сдач, то это тоже не проблема. Поскольку получение «стандартизованных» match point в сдаче имеет нормальное распределение с параметрами $\mathbb{N}(0,1)$, то за k сдач вероятностное распределение будет иметь параметры $\mathbb{N}(0, \sqrt{k})$. Это означает, что стандартизованные match point можно переводить в проценты с учетом количества сыгранных сдач⁴. Сами проценты можно интерпретировать, как вероятность выигрыша, оцененную через k сдач.

⁴ Расчет интегрального нормального распределения зашит практически во все компьютерные программы.

На практике, вычисляя интегральное нормальное распределение с параметрами $N(0, \sqrt{k})$, можно округлять до 2-х знаков после запятой, что дополнительно сглаживает полученные результаты. Так можно получать значение victory point.

Проблема «выбросов»

При использовании указанного алгоритма остро встает вопрос о «выбросах», когда в протоколе оказывается «необычная» запись «необычного» контракта. Предложенный алгоритм, как и алгоритм «Игры на среднее», очень чувствителен к таким записям.

В «Игре на среднее» эта проблема решается через вычисление урезанного среднего, когда из протокола выбрасываются записи 25% наименьших и 25% наибольших.

Однако для предложенного алгоритма оценки стоимости сдачи он не подходит – слишком много теряется информации для расчета s .

Для решения этой проблемы можно устанавливать предельное значение mp , которое может получить пара за сдачу. Например, это могут быть значения:

mp	$N(0,1)$
1.645	95.00%
2.000	97.72%
2.326	99.00%
3.000	99.87%