LUISS

Department
of Business and Management

Course of Social Network Analysis

Regression Analysis of Contract Bridge Hand Evaluation and its Application on Bidding

Prof. Marco Scarsini

SUPERVISOR

CO-SUPERVISOR

Fangqing Yuan, ID No. 232231

CANDIDATE

Academic Year 2020/2021

Table of contents:

Summary:

The thesis addresses the Hands Evaluation problem in Contract Bridge, particularly the value assignment of High Cards for no-trump contracts. We develop two new point-count hand evaluation methods for No-Trump contracts based on regression analysis of randomly generated deals. The study compare new methods with traditional High-Card-Point (HCP) method and analyze the advantage of the new methods.

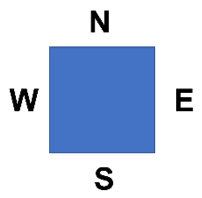
Key words:

Contract Bridge, Hand Evaluation, Point-Count, Bidding, Regression Analysis

1. Introduction

1.1 Contract Bridge

Contract Bridge, or shortly, **Bridge**, is a card game using a standard 52-card deck (each card has a color and a number, the deck of cards is a cartesian product of colors { **Clubs (♣), Diamonds (♦), Hearts (♥), Spades (♠)**} and numbers {2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A}). In a single game, it is played by four players in two competitive pairs, when partners sit opposite to each other around a four-side table. Four players are referred as **North(N), South(S), East(E), West(W)**.



Picture 1, Bridge players and their positions

Bridge games consist of an agreed number of **Deals**, and each deal has four phases:

a) Dealing

Each player is dealt 13 (exactly one fourth of 52) cards from the deck. The set of all cards one player has in one deal is called a **Hand**.

b) Bidding

The players bid in turn clockwise in an auction to decide the **Contract** of the deal. Players rotate to be the starter (opener) of bidding.

A contract consists of two attributes: **Level** and **Trump**. A level ***n*** (an integer from 1 to 7) means that the partnership who bids the contract should win at least $n+6$ **Tricks** (to be introduced in next phase) to **Make** it. The trump could be one of the four colors **Clubs (♣), Diamonds (♦), Hearts (♥), or Spades (♠), or No-Trump (NT)**.

All possible contracts are shown in the picture below (from 1**♣ to 7NT)**, while players also have option to **Pass, Double (X)**, and **Re-double (XX)** during the auction. Three consecutive Pass's after the last contract bidding marks the end of Bidding phase, while the last bidden contract is to be played by the pair who bid it.

Picture 2, All possible biddings in bridge game.

(Retrieved from https://www.funbridge.com/how-to-play-bridge)

c) Playing

The contract is going to be played by a pair of players, and the one who has bidden the contract trump (could be either a color or no-trump) first becomes the **Declarer**, the partner of declarer becomes the **Dummy**, and the **Left-Hand-Opponent**(LHO, the clockwise next player on the table) of declarer becomes the one who **Lead**s, which means he/she plays the first card (**The Lead**) of the playing phase. After the Lead is played, the dummy is obliged display all his/her cards on the table to everyone, and the declarer will control the plays of the dummy in this playing phase. The opponent of declarer and dummy is called **Defense** or **Defender**.

Playing phase consists of 13 **Rounds**. During each round each player plays 1 card without replacement (Any played card should be displaced from players' hand and never be played again in the deal) in a clockwise sequence. The Color (**♣, ♦, ♥, ♠)** of each round is the color of the card played by the first player. Each of the following three players of each round must play the same **color** if one has at least one card of the color in his/her remaining cards, otherwise one can **Discard** or **Ruff**:

Discard: play a card whose color is neither the color of the round nor the color of the contract trump. (If the contract is no-trump, there is no trump color)

Ruff: play a card of trump color, which would potentially win the round. (In no-trump contract there is no Ruff)

In each round, the pair of the winning player wins the round, and it is recorded as one **Trick** won by the pair. The player who won the trick starts to play the next round. The winner is decided by comparing the four played cards in a round as following:

i) Compare the color: The trump color is superior to the color of the round, which is superior to other colors.

ii) Compare the number, the numbers from the lowest to the highest are 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A

For example:

1) If the contract is 1♣**,** and **♦Q, ♦K,** ♣4, ♣A are played in order in one round. ♣A should win this round since ♣A > ♣4 > **♦K** > **♦Q**
2) If the contract is 3NT**,** and **♦4, ♦10,** ♦Q, ♣9 are played in order in one round. **♦Q** should win this round since ♦Q > ♦10 > **♦4** > **♣9**

d) Scoring

After 13 rounds, players count how many tricks are won by each partnership. There are certain rules to decide how much score each pair would be given. Making the contract by winning exact number of tricks is called "**Just Make**", fulfilling extra tricks is called "**Plus** by n tricks", and failed to make the contract by n tricks is called "**Down** by n tricks". The scoring is slightly different when the pair is **Vulnerable** or **non-Vulnerable**, each appears in 50% of the deals. (In vulnerable condition, rewards for making a contract tend to be higher than those when non-vulnerable, while punishment for not making a contract tend to be more severe than those when non-vulnerable)

In bridge competition the game winner is not decided by the absolute amount of scores gained by a pair of players. The main idea is to let different players play the same deal and see who can gain more score.

Here is an example of a bidding phase in a bridge game:

N	E	S	W
PASS	PASS	1NT	PASS
3NT	PASS	PASS	PASS

Picture 3, a bidding example.

North started with a PASS, the final contract would be 3NT since there were 3 consecutive PASSES after N bids 3NT. Although N bids 3NT, S was the player who bid NT first in their partnership, thus S would be the declarer playing this contract. W, as the left-hand opponent of S, would lead the playing phase, playing the first card of the first round. The contract 3NT means that N and S should win at least 3(level of contract)+6=9 tricks out of 13 to make the contract, and during the play, there is no trump color.

1.2 Bridge Bidding System

To reach the best contract in the auction, players have developed **Conventions** which assign meanings to each bidding to facilitate partners exchange information of their cards in the bidding phase. A set of bidding conventions constitute a bidding system.

The information usually includes but is not limited to **Distribution** and **Strength**. The process of assessing these attributes is called **Hand Evaluation**.

a) Distribution/Shape

The shape means how many cards one hand has in each four colors. It is usually marked as a decreasing sequence of numbers, such as 4-4-3-2, 5-4-3-1, 6-3-2-2. In this case, 4-4-3-2 means one has four, four, three, two cards in each color, but not necessarily in a specific order of four colors. Since one hand has 13 cards, the four numbers must sum up to 13.

Here are some terms that will be used later:

Suit, a set of all card(s) one owns in one color. For example, “a 5-card Spade suit”, which means all the five cards one hand owns.

Balanced Hand, a hand which has one of the following distributions: 4-3-3-3, 4-4-3-2, 5-3-3-2. It is called balance since the cards are approximately distributed among four colors.

Singleton, a one-card suit

Void, a zero-card suit. It means there is no card in a certain color in one hand.

Semi-Balanced Hand, a hand with no Singleton, no Void, no 7-card-suit. The possible shapes are: 4-3-3-3, 4-4-3-2, 5-3-3-2, 5-4-2-2, 6-3-2-2.

b) Strength

As previously shown, in each color, there are cards superior to others. Therefore, to assess the potential to win tricks of a hand, players care about how many **High Cards** , such as A, K, Q, J, one has.

Many bidding systems use point-count methods to do hand evaluation. One of the most widely used method is the **4-3-2-1** point-count. The system introduces the idea of **High-Card-Point (HCP)** and assigns the HCP to high cards as follows:

A = 4 HCP

K = 3 HCP

Q = 2 HCP

J = 1 HCP

In a 52-card deck, there are 4 A, 4 K, 4 Q, and 4 J, and therefore we have 4x(4+3+2+1)=40 HCP in total for a bridge game deal.

The HCP method was first introduced by Bryant McCampbell[1] in 1915, derived from the idea in another card game Auction Pitch. It is also referred as Milton Work count or Work-Goren count since it was popularized by Work and Goren[2].

The players may therefore count the HCP of their hands and bid according to their conventions.

Picture 4, an example hand

For example, in ACBL Standard American Yellow Card system[3], opening bidding 1NT means 15~17 HCP, Balanced hand. In the hand shown above, there are 2 A, 2 K, and 1 Q, which combine to be 16 HCP in total, and it has a distribution of 4-4-3-2, which is balanced. Therefore, the play who holds this hand should bid 1NT at the first place.

1.3 No-Trump Contract

The scoring for no-trump contracts are calculated by following rules:

Let the contract be *x*NT, and the declarer side win *t* tricks at the end of playing phase. $x \in \{1,2,3,4,5,6,7\}, t \in \{1,2,3,4,5,6,7,8,9,10,11,12,13\}$.

i) If the contract is made ($t \geq x + 6$), the **Base Score** is $40 + 30 \times (t - x - 6)$.

If $40 + 30 \times (x - 1) < 100$, equivalently $x < 3$, the making **Bonus** is 50.

If $40 + 30 \times (x - 1) \geq 100$, equivalently $x \geq 3$, the Game **Bonus** is 300 if non-vulnerable, 500 if vulnerable.

If $x = 6$, the **Small Slam Bonus** is 500 if non-vulnerable, 750 if vulnerable.

If $x = 7$, the **Big Slam Bonus** is 1000 if non-vulnerable, 1500 if vulnerable.

The **final score** for making the contract = $Base\ Score + Bonus + Small\ Slam\ Bonus + Big\ Slam\ Bonus$

ii) If the contract is not made ($t < x + 6$), the score (**punishment**) is:

$(-50) \times (x + 6 - t)$ if non-Vulnerable,

[1] McCampbell, B (1915), *Auction tactics*, P.26, Dodd, Mead and Company
[2] Goren,C.H. (1961), *Point count bidding in contract bridge*, P.11, Simon and Schuster
[3] American Contract Bridge League (2006), *ACBL SAYC SYSTEM BOOKLET*, ACBL

$(-100) \times (x + 6 - t)$ if Vulnerable.

Example:

If N and S played **3NT** when vulnerable, winning 9 tricks (x=3, t=9), the Base Score is 100, Bonus is 500, and there is no Slam Bonus. The score for N and S would be **600**.

If N and S played **1NT** when non-vulnerable, also winning 9 tricks (x=1, t=9), the Base Score is 100, Bonus is 50, and there is no Slam Bonus. The score for N and S would be **150**.

If N and S played **3NT** when vulnerable, winning 8 tricks (x=3, t=8), there will be no Base Score and Bonus but negative score punishment. The score for N and S would be **-100**.

2. Existing Literature

There have been academic studies regarding bridge since its birth. The most renowned one is probably Émile Borel and André Chéron's ***Théorie mathématique du bridge à la portée de tous*** (English: Mathematical Theories of Bridge for Everyone) first published in 1940. The book conducted a comprehensive statistical study on bridge regarding cards shuffling, a priori probabilities of hands, shapes, and occurrence of high cards and suits, and a posteriori probabilities of those after biddings, which build a solid foundation for future bridge players and researchers. It also proposed an idea to assign "valeur défensive" (defensive value) to certain high cards combinations by computing how many tricks the cards would win on average using large data sample.

The idea of assigning value to cards is similar to the hand evaluation we mentioned. After the introduction of HCP, players started to question its accuracy. For example, the A and 10 were underestimated while K, Q, J were overly valued and there were computer analysis supporting it **(Bergen, 2002)**. Researchers began to develop more accurate value-assignment for high cards **(Woolever, 2000)**. In *Evolving a Bridge Bidding System*, Woolever claimed that appropriate high cards values could be calculated using double-dummy results (double-dummy refers to the situation where each player is aware of what specific cards each other player has. This perfect information is never available in real games but could provide game result when all players could play optimally). Woolever designed a program called WEBB for bridge bidding system using Mathew Ginsberg's library[4] of double dummy deals to compute relevant values. He gave an example of this method: "An ace, on average, produces 1.391 tricks, so it has an HCP value of 1.391 x 40/13 = 4.280.". He also pointed out that using double dummy data may have potential problems since the situation is different from real life game and players may not find the optimal playing as they could do in double-dummy situation. He claimed that "This difference is minor, however, and will not adversely affect the results" without justifying it.

[4] http://hp.vector.co.jp/authors/VA051022/ginsberg/GetLibrary.HTML Matthew Ginsberg's library of double dummy results

Further attempts to invent hand evaluation methods were made, and one the most famous point-count systems among bridge players is Zar Points. In Zar Petkov's ***Zar Points – Aggressive Bidding Hand Evaluation*** published in 2003, Petkov analyzed "hundreds and hundreds of "aggressive" game contracts bid by world-class experts" and invented a new system called Zar Points. In Zar Points, not only high cards are assigned values, other distribution attributes like shapes, long suits, singleton, and voids are considered for this point-count system. Petkov also provided a method to compare the performance of Zar Points with HCP. He started with generating massive amounts of deals of which a certain contract could be made in double-dummy scenario, then he checked if Zar Points bidding system and HCP bidding system could bid this contract or not, and therefore calculated the success bidding rates of two systems.

3. Research Question: Hand Evaluation in No-Trump Contracts

The study focuses on no-trump contracts for the following reasons:

1) In no-trump contracts there is no Ruffing. Therefore, the possession of high cards is one of the most important parts of hand evaluation.

2) A no-trump contract is reached usually when both partners have balanced or semi-balanced hands. Semi-Balanced hands make up to 63.83% probability of occurrence of all possible distributions (Borel et al, 1940).

3) 28.45% of contracts in real life are No-Trump contracts[5].

In HCP system, there are in total 40 HCP in four hands and thus 10 HCP for each hand in average. Usually it is agreed that a pair with 21~24 HCP in total should play 1NT or 2NT, a pair with 25 or more HCP should play 3NT, 33 HCP or more should play 6NT, and 37 HCP or more for 7NT. We will demonstrate statistically the reasoning later.

This aims at solving a problem that is commonly faced by bridge players: "should we bid 3NT or stay at 2NT?". As previously shown, the score for making 3NT is much higher than making 2NT. (while vulnerable, making 3NT gains 600 scores, while reaching 9 tricks in 2NT also provides 150 scores). This significant difference calls for accurate techniques to help players make correct decisions.

In our research, we are going to develop a new point-count hand evaluation method specifically for no-trump contracts bidding on level 2 and 3 and test its performance comparing to HCP. The new

[5] http://www.bridgetoernooi.com/index.php/home/stats a study on 130000 bridge deals from data of vugraph project (http://www.sarantakos.com/bridge/vugraph.html)

method would be useful for bridge players to improve their performance and serve as a source for developing bridge bidding programs.

4. Regression Analysis

4.1 Analysis pre-design

We decide to follow Woolever's idea of using double-dummy results to evaluate high cards. However, we plan not to use Matthew Ginsberg's data but aim to randomly generate a massive amount of bridge deals under certain distribution and strength restriction. For each generated deal, a no-trump contract could be assigned to compute the double-dummy result and score.

We will record the number of high cards that declarer and dummy had and use them as regressors while using the number of tricks they won as the response variable. The coefficient of each high card would indicate how many tricks it can produce on average, and it could be used to assign points to high cards.

The study performs ordinary least squares (**OLS**) linear regressions on the data gained from generated deals. As suggested by Zar Points and other experts, we decide to add other attributes regarding the shapes of hands. As we only consider semi-balanced hands in the study, we decide to record if the hands had a 6-card-suit or not in each deal, as experience tells that a long-suit might benefit a no-trump contract. After generating several regression models, we will evaluate these models and pick one as the guideline for our new point-count system.

4.2 Tools

The tools we use is ***redeal***[6], a bridge-deal-generator coded in Python3 which randomly outputs bridge deals under certain conditions that users specify. *Redeal* is also integrated with a **double-dummy-solver** (DDS) based on Bo Haglund's DDS[7] 2.9.0. A **DDS** is a function that takes a bridge deal's four hands and a given contract as input, and outputs the contract result under optimum playing of declarer and defenders (It computes the result under double-dummy situation). For regression we used function **OLS** from Python package **statsmodels**.api, and we use **scipy** and **R** as tool for statistical analysis.

4.3 Analysis process

The following describes the procedure we used to generate regression models using Python:

i) Randomly generate 100,000 (1e+5) bridge deals using *redeal*. The deals must satisfy the following conditions:
   1. Both N and S have semi-balanced hands.
   2. N and S have 22~28 HCP in total

[6] redeal, https://github.com/anntzer/redeal

[7] Bo Halund's DDS and its algorithm, http://privat.bahnhof.se/wb758135/bridge/dll.html

ii) For each deal, record how many tricks N and S would get in a No-Trump contract using the DDS. Name the variable as *tricks*.

The data gained should be a vector $tricks$,

$$for\ i\ in\ \{1, 2, \ldots 100000\}, tricks_i$$
$$= the\ number\ of\ tricks\ gained\ by\ N\ and\ S\ in\ the\ ith\ deal$$

$0 \leq tricks_i \leq 13$.

iii) For each deal, record separately the number of A, K, Q, J, 10, 9 that N and S have in total. Name the variables as *As*, *Ks*, *Qs*, *Js*, *Ts*, *Ns*. (T stands for 10, N stands for 9)

The data gained should be 6 vectors As, Ks, Qs, Js, Ts, Ns.

Take *As* as an example:

$$for\ i\ in\ \{1, 2, \ldots 100000\}, As_i$$
$$= the\ number\ of\ A\ that\ N\ and\ S\ have\ in\ total\ in\ the\ ith\ deal$$

$0 \leq As_i \leq 4$.

The other variables should have the same characteristics.

iv) Create a new variable *sixs*, and, for each deal. Record how many six-card suits N and S have in total.

The data gained should be a vector $sixs$,

$$for\ i\ in\ \{1,2, \ldots 100000\}, sixs_i$$
$$= the\ number\ of\ 6 - card$$
$$- suits\ that\ N\ and\ S\ have\ in\ total\ in\ the\ ith\ deal$$

$0 \leq sixs_i \leq 2$.

v) Create 3 OLS Linear Regression models:

1) Model_1: (*As*, *Ks*, *Qs*, *Js*, *Ts*, *Ns*) as regressor, *tricks* as response variable.

$$tricks\ =\ \beta_0 + \beta_1 \times As + \beta_2 \times Ks + \beta_3 \times Qs + \beta_4 \times Js + \beta_5 \times Ts + \beta_6 \times Ns$$

2) Model_2: (*As*, *Ks*, *Qs*, *Js*, *Ts*) as regressor, *tricks* as response variable.

$$tricks\ =\ \beta_0 + \beta_1 \times As + \beta_2 \times Ks + \beta_3 \times Qs + \beta_4 \times Js + \beta_5 \times Ts$$

3) Model_3: (*As*, *Ks*, *Qs*, *Js*, *Ts*, *sixs*) as regressor, *tricks* as response variable.

$$tricks\ =\ \beta_0 + \beta_1 \times As + \beta_2 \times Ks + \beta_3 \times Qs + \beta_4 \times Js + \beta_5 \times Ts + \beta_6 \times sixs$$

4.4 Models' generation and evaluation:

Here are the summaries of three generated models: (the coefficients xi shown in pictures are indicated as estimated coefficients of βi in formulas above.

OLS Regression Results

Dep. Variable:	y	R-squared:	0.426
Model:	OLS	Adj. R-squared:	0.426
Method:	Least Squares	F-statistic:	1.238e+04
Prob (F-statistic):	0.00	Log-Likelihood:	-1.5663e+05
No. Observations:	100000	AIC:	3.133e+05
Df Residuals:	99993	BIC:	3.133e+05
Df Model:	6		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-4.2407	0.052	-82.235	0.000	-4.342	-4.140
x1 (As)	2.1968	0.008	266.648	0.000	2.181	2.213
x2 (Ks)	1.4410	0.007	214.942	0.000	1.428	1.454
x3 (Qs)	0.8240	0.005	155.101	0.000	0.814	0.834
x4 (Js)	0.4316	0.004	101.364	0.000	0.423	0.440
x5 (Ts)	0.1937	0.004	50.398	0.000	0.186	0.201
x6 (Ns)	0.0695	0.004	18.060	0.000	0.062	0.077

Omnibus:	2520.015	Durbin-Watson:	2.001
Prob(Omnibus):	0.000	Jarque-Bera (JB):	3627.770
Skew:	-0.279	Prob(JB):	0.00
Kurtosis:	3.748	Cond. No.	78.7

Picture 5, Regression Model_1

OLS Regression Results

Dep. Variable:	y	R-squared:	0.424
Model:	OLS	Adj. R-squared:	0.424
Method:	Least Squares	F-statistic:	1.474e+04
Prob (F-statistic):	0.00	Log-Likelihood:	-1.5679e+05
No. Observations:	100000	AIC:	3.136e+05
Df Residuals:	99994	BIC:	3.137e+05
Df Model:	5		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-4.0164	0.050	-80.117	0.000	-4.115	-3.918
x1 (As)	2.1885	0.008	265.624	0.000	2.172	2.205
x2 (Ks)	1.4326	0.007	213.858	0.000	1.419	1.446
x3 (Qs)	0.8154	0.005	153.850	0.000	0.805	0.826
x4 (Js)	0.4230	0.004	99.811	0.000	0.415	0.431
x5 (Ts)	0.1848	0.004	48.405	0.000	0.177	0.192

Omnibus:	2437.423	Durbin-Watson:	2.000
Prob(Omnibus):	0.000	Jarque-Bera (JB):	3485.612
Skew:	-0.274	Prob(JB):	0.00
Kurtosis:	3.732	Cond. No.	71.9

Picture 6, Regression Model_2

OLS Regression Results

Dep. Variable:	y	R-squared:	0.426
Model:	OLS	Adj. R-squared:	0.426
Method:	Least Squares	F-statistic:	1.238e+04
Prob (F-statistic):	0.00	Log-Likelihood:	-1.5663e+05
No. Observations:	100000	AIC:	3.133e+05
Df Residuals:	99993	BIC:	3.133e+05
Df Model:	6		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-4.0499	0.050	-80.859	0.000	-4.148	-3.952
x1 (As)	2.1896	0.008	266.178	0.000	2.173	2.206
x2 (Ks)	1.4331	0.007	214.286	0.000	1.420	1.446
x3 (Qs)	0.8157	0.005	154.166	0.000	0.805	0.826
x4 (Js)	0.4231	0.004	99.994	0.000	0.415	0.431
x5 (Ts)	0.1851	0.004	48.542	0.000	0.178	0.193
x6 (sixs)	0.1679	0.009	18.017	0.000	0.150	0.186

Omnibus:	2815.278	Durbin-Watson:	1.999
Prob(Omnibus):	0.000	Jarque-Bera (JB):	4013.246
Skew:	-0.307	Prob(JB):	0.00
Kurtosis:	3.766	Cond. No.	72.0

Picture 7, Regression Model_3

1) In Model_1, the p-value for x6, the coefficient of Ns, is < 0.05, although the coefficient is low compared to other variables. Therefore we considered it as a valid acceptable model while we can further test the difference of it with Model_2.
2) Then we focus on Model_2, in which we kept Ns out of the model compared to Model_1. It has a high R-square, and all p-values for coefficients are low, which made it an acceptable model.
3) In Model_3, the added attribute *sixs* has a significant coefficient. It indicates that a 6-card-suit may have value that is close to the value of 10. It would be helpful if we consider distribution in our new point-count system.

5. Develop New Point-Count Method using generated model

We interpret the coefficients in this way:

*The coefficient **xn** for card **N** indicates that, on average, a card **N** produces **xn** tricks.*

Woolever assigned points to high cards based on this reasoning:

"If a hand with one K tends to produce one more trick than a hand with no K, a K could be given a weight of 1/13 the total of all cards." (**Woolever, 2000**)

His method assigns 1.391 x 40/13 = 4.280 points to A due to 1.391 tricks generated by an A on average. However, this method doesn't generate points of cards that sum up to 10 (40 in whole deck) as HCP does. Therefore, we decide to re-scale the coefficients so that they sum up to 10 to facilitate future application:

$$given\ all\ n\ coefficients\ x_i, we\ have\ n\ values\ p_i = x_i \times 10/\sum x_i$$

For programming reasons, we decided to only convert Model_1 and Model_2 into point-count methods.

In Model_1, The coefficients for A, K, Q, J, T(10), N(9) are 2.1968, 1.4410, 0.8240, 0.4316, 0.1937, 0.0695. By re-scaling the coefficients, we are able to assign value points to each high cards as the following: (we name the unit for this new point as **Brand New Point, BNP**)

A = 4.26 BNP

K = 2.80 BNP

Q = 1.60 BNP

J = 0.84 BNP

T = 0.37 BNP

N = 0.13 BNP

In Model_2, The coefficients for A, K, Q, J, T(10) are 2.1885, 1.4326, 0.8154, 0.4230, 0.1848. By re-scaling the coefficients, we are able to assign value points to each high cards as the following: (we name the unit for this new point as **Whole New Point, WNP**)

A = 4.34 WNP

K = 2.84 WNP

Q = 1.62 WNP

J = 0.84 WNP

T = 0.36 WNP

6. Application of BNP and WNP on bidding and its performances

6.1 Assumptions on bidding application

To simplify the test, we ask a simple question for each deal: 2NT, or 3NT?

As we previously discussed, traditionally bridge players agree that there is a **threshold** for total amount of HCP for a pair to decide if they should bid 3NT (it tells the players to bid 3NT only if they have at least this amount of HCP in total). Usually this threshold is 25 HCP or 24 HCP. To testify this statement, we design this method using two-sample t-test to test that, given a certain HCP in two hands, whether they should bid 3NT or not. This method would be also used on our BNP and WNP:

i) Randomly generate *n* deals using *redeal*, in which N and S have semi-balanced hands and total HCP of *h*. (*h* is the threshold we want to test, in case of HCP, it could be 24 or 25)

ii) Using DDS, calculate the score N and S would get if they played 3NT for each deal

iii) Using DDS, calculate the score N and S would get if they played 2NT for each deal

iv) We therefore get two sets of data containing scores of two contracts for each deal

The data gained should be two vectors $score_{3NT}, score_{2NT}$, for each vector it satisfies:

$$for\ i\ in\ \{1, 2, \ldots n\}, score_i = the\ score\ gained\ by\ N\ and\ S\ in\ the\ ith\ deal$$

$score_i \in R$.

v) Perform a **Welch**'s independent samples t-test[8] on these two samples $tricks_{3NT}, tricks_{2NT}$, and decide if we should reject the null hypothesis that "scores gained by playing 3NT is not greater than scores gained by 2NT."

vi) If we reject the null hypothesis of t-test, it means that bidding 3NT is worthier than bidding 2NT if the pair has *h* HCP in total. If we don't reject the null hypothesis, it means bidding 3NT is not a good idea for *h* HCP. (We choose 95% as our confidence level)

vii) Find the lowest *h* that rejects the null hypothesis, and it would be the threshold that we were looking for.

[8] Welch (1947), The Generalization of `Student's' Problem when Several Different Population Variances are Involved

viii) Perform this method in both vulnerable and non-vulnerable conditions. Find thresholds for both conditions. (they might be different)

Test results by using n=1000 deals each test for **HCP**:

For vulnerable conditions:

p-value when HCP=24 is: = 6.75e-06

p-value when HCP=23 is: = 0.9863

The p-values show that bidding 3NT is worthy when HCP=24 but not worthy when HCP=23,

Therefore, we set the threshold for bidding 3NT as 24 HCP when vulnerable.

For non-vulnerable conditions:

p-value when HCP=25 is: = 2.2e-16

p-value when HCP=24 is: = 0.01319

p-value when HCP=23 is: = 0.9997

The p-values show that bidding 3NT is worthy when HCP=24 but not worthy when HCP=23,

Therefore, we set the threshold for bidding 3NT as 24 HCP when non-vulnerable.

However, for our BNP and WNP, such threshold is more difficult to determine than for HCP, since (4, 3, 2, 1) has a maximum common divisor 1 while the points values of BNP and WNP are with decimals.

Therefore, instead of giving a certain value h, we decide to set an interval $hn = (ha, hb)$ when performing each test. Based on the model we obtained, we can assume that number of tricks won is proportional to the points. So we start from guessing a small interval, do its t-test, and move to test other intervals based on the previous test result.

Test results by using n=1000 deals each test for **BNP**:

For vulnerable conditions:

p-value when BCP∈[23.6, 23.7] is: = 0.009579

p-value when BCP∈[23.55, 23.65] is: = 0.04621

p-value when BCP∈[23.5, 23.6] is: = 0.2205

The p-values show that bidding 3NT is worthy when BNP∈[23.55, 23.65] but not worthy when HCP∈[23.5, 23.6]. Setting 23.55 as threshold would be too bold.

Therefore, we approximate the threshold for bidding 3NT as 23.6 BNP when vulnerable.

For non-vulnerable conditions:

p-value when BCP∈[23.9, 24.0] is: = 0.002091

p-value when BCP∈[23.85, 23.95] is: = 0.07064

p-value when BCP∈[23.8, 23.9] is: = 0.1058

The p-values show that bidding 3NT is worthy when BNP∈[23.9, 24.0] but not worthy when HCP∈[23.85, 23.95]. Setting 23.9 as threshold might be low.

Therefore, we approximate the threshold for bidding 3NT as 23.95 BNP when non-vulnerable.

Test results by using *n*=1000 deals each test for **WNP**:

For vulnerable conditions:

p-value when WCP∈[23.7, 23.8] is: = 0.000133

p-value when WCP∈[23.65, 23.75] is: = 0.0002949

p-value when WCP∈[23.6, 23.7] is: = 0.03552

p-value when WCP∈[23.5, 23.6] is: = 0.7926

The p-values show that bidding 3NT is worthy when BNP∈[23.6, 23.7] and when BNP∈[23.65, 23.75] but not worthy when HCP∈[23.5, 23.6]. Setting 23.6 as threshold would be too low, while 23.65 would probably guarantee a good result.

Therefore, we approximate the threshold for bidding 3NT as 23.65 WNP when vulnerable.

For non-vulnerable conditions:

p-value when WCP∈[24.0, 24.1] is: = 3.417e-05

p-value when WCP∈[23.95, 24.05] is: = 0.06379

p-value when WCP∈[23.9, 24.0] is: = 0.589

p-value when WCP∈[23.8, 23.9] is: = 0.913

The p-values show that bidding 3NT is worthy when BNP∈[24.0, 24.1] but not worthy when HCP∈[23.95, 24.05]. Setting 24.0 as threshold would be slightly low, while 24.0 would probably guarantee a good result.

Therefore, we approximate the threshold for bidding 3NT as 24.05 WNP when non-vulnerable.

System\Condition	Vulnerable	Non-vulnerable
HCP	24	24
BNP	23.6	23.95
WNP	23.65	24.05

Table 1, threshold for bidding 3NT in different systems and conditions

6.2 Performance test and analysis

After introducing BNP, WNP and their corresponding threshold for 3NT, we have to test if we could use them to get better game result than using traditional HCP.

We test the performance of HCP, BNP, and WNP in the following ways:

i) Randomly generate *n* deals using *redeal*, in which N and S have semi-balanced hands, with 23~24 HCP in total.

ii) Compute the total HCP, BNP, and WNP in two hands.

iii) Based on the amount of points and vulnerable/non-vulnerable condition, decide if we should bid 3NT or 2NT in this deal according to the thresholds we calculated in 6.1

$$\text{Given: } method \in \{\text{HCP}, \text{BNP}, \text{WNP}\}, points \in [0,40],$$

$$condition \in \{\text{vulnerable}, \text{non} - \text{vulnerable}\}, \text{use results in 6.1 to find the } threshold.$$

$$If\ points \geq threshold: bid\ and\ play\ 3NT$$

$$If\ points\ < threshold: bid\ and\ play\ 2NT$$

iv) Using DDS, calculate the score N and S would get if they played 3NT or 2NT for each deal.

v) Assign and record the scores of using HCP, BNP, WNP accordingly.

The data gained should be three vectors $score_{HCP}, score_{BNP}, score_{WNP}$, for each vector it satisfies:

$$for\ i\ in\ \{1,2, \dots n\}, and\ method\ \in \{HCP, BNP, WNP\}:$$

$$score_{method,i}$$
$$= score\ gained\ by\ N\ and\ S\ of\ playing\ the\ contract\ decided\ by\ using\ the\ method\ indicated$$

$score_i \in R$.

vi) Perform Welch's independent samples t-test separately on $score_{HCP}\ and\ score_{BNP}, score_{HCP}\ and\ score_{WNP}$. Decide if we should reject the null hypothesis that

1) "Scores gained by using NBP is not greater than scores gained by using HCP."
2) "Scores gained by using WBP is not greater than scores gained by using HCP."

t-test Results:

```
> t.test(bnp, hcp, var.equal = FALSE, alternative = 'greater')

        Welch Two Sample t-test

data:  bnp and hcp
t = 1.4913, df = 1994.2, p-value = 0.06802
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -1.80381      Inf
sample estimates:
mean of x mean of y
    96.88     79.45
```

Picture 8, t-test result for BNP and HCP (Vulnerable)

```
> t.test(wnp, hcp, var.equal = FALSE, alternative = 'greater')

        Welch Two Sample t-test

data:  wnp and hcp
t = 1.5122, df = 1992.5, p-value = 0.06532
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -1.55273      Inf
sample estimates:
mean of x mean of y
    97.05     79.45
```

Picture 9, t-test result for WNP and HCP (Vulnerable)

```
> t.test(bnp, hcp, var.equal = FALSE, alternative = 'greater')

        Welch Two Sample t-test

data:  bnp and hcp
t = 1.9069, df = 1927.6, p-value = 0.02834
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 1.827555      Inf
sample estimates:
mean of x mean of y
   106.31     92.97
```

Picture 10, t-test result for BNP and HCP (Non-vulnerable)

```
> t.test(wnp, hcp, var.equal = FALSE, alternative = 'greater')

        Welch Two Sample t-test

data:  wnp and hcp
t = 1.5989, df = 1892.5, p-value = 0.05501
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -0.3211786       Inf
sample estimates:
mean of x mean of y
   103.95     92.97
```

Picture 11, t-test result for WNP and HCP (Non-vulnerable)

As the results shown above, in both vulnerable and non-vulnerable conditions, the mean values of BNP and WNP are higher than HCP. The p-values for testing BNP and HCP in non-vulnerable situation is < 0.05, while others are just slightly higher than 0.05.

We can say that using BNP or WNP would gain score on average, and the difference is statistically significant when we are using BNP in non-vulnerable condition.

More significant results could be obtained if we could approximate the thresholds for BNP and WNP more accurately.

7. Conclusions

We proposed the use of regression analysis to assign more accurate values for high cards in no-trump contracts bidding. Two new point-count systems (BNP and WNP) were designed to make 3NT bidding decisions. While the performance tests showed that BNP and WNP perform better than traditional HCP, the difference are not all statistically significant. We have proposed the idea to include information of long suits into the model but not further studied it due to several constraints. Some other aspects, as previous studies indicated, such as length of trump suit, various combinations of high cards, existence of singleton and void, could be further considered and be used to develop finer hand evaluation methods for not only no-trump contracts we discussed. Those new methods could be studied and applied into real life bridge competitions to improve players' performance, and even serve as a foundation to develop bridge bidding programs.

8. Bibliography

1] McCampbell, B (1915), Auction tactics, Dodd, Mead and Company

2] Goren,Charles. (1961), Point count bidding in contract bridge, Simon and Schuster

3] Goren, Charles (1949). Point Count Bidding in Contract Bridge, Simon and Schuster

4] American Contract Bridge League (2006), ACBL SAYC SYSTEM BOOKLET, ACBL

5] Émile, Borel; André, Chéron (1940). Théorie Mathématique du Bridge a la portée de tous, Paris: Libr. Gauthier-Villars

6] Bergen, Marty (2002). Hand Evaluation: Points Schmoints, Bergen Books

7] Woolever, J.R. (2000). Evolving a Bridge Bidding System, Massachusetts Institute of Technology

8] Petkov, Zar (2003). Zar Points – Aggressive Bidding Hand Evaluation.

9] Haglund, Bo (2014). Search Algorithms for a Bridge Double Dummy Solver

10] Welch (1947), The Generalization of `Student's' Problem when Several Different Population Variances are Involved