Skill in Backgammon: Cubeful vs Cubeless

Tilemachos Zoidis, June 2025
bloarootz@yahoo.gr

Abstract. Does the doubling cube make backgammon more skillful? And is the answer the same in both money and match play? This paper presents GNUbg rollouts between unequally skilled players, which show that use of the doubling cube does not favor the better player in either case.

Keywords: Backgammon, Cubeful, Cubeless, Skill, ELO

Contents

1. Introduction

Luck is arguably the most common thing that backgammon players complain about. The doubling cube, a rather recent invention compared to the long history of backgammon, introduces a new element of skill and has therefore been touted as a way to reduce luck. But is this actually true? And if so, is it true in both money and match play? To answer this question, we used GNUbg rollouts between unequally skilled players, which show that use of the doubling cube does not favor the better player in either case. We also examine the implications of these data on the ELO system.

2. Money Play

We will compare 3 types of money games:

- cubeful games with the Jacoby rule in effect (cubeful games),
- cubeless games with gammons and backgammons (cubeless games) and
- cubeless games without backgammons or gammons (DMP games).

In order to compare the skill in these formats, we need 2 things for each one:

- the equity E of the better player (the expected difference in points per game) and
- the expected value V of a game (the average PPG of the winner, assuming optimal play).

2.1. Expected Value of Money Games

For cubeless games, Tom Keith has rolled out every opening roll 46,656 times using GNUbg 2ply and reported a **gammon rate of 27.62%** and a **backgammon rate of 1.22%** [1]. For cubeful games, we used GNUbg 2ply to roll out 19,440 games. These data yield the following results:

$$
\begin{aligned}
V_{DMP} &= 1\,ppg \\
V_{CL} &= 1.2884\,ppg \\
V_{CF} &= 2.3856\,ppg \; [SE = 0.0123\,ppg]
\end{aligned}
$$

2.2. Equity of the Better Player

Here **1,399,680 games** were rolled out for each format, with GNUbg playing one side at 1ply and the other at 0ply. These games were rolled out without Variance Reduction (VR). Because of the way it works, VR would actually skew the results. VR works by subtracting the sum of the equity differences between 2 consecutive plies from the final result. This is often interpreted as canceling out the estimated luck but it's equivalent to think about it as using subsequent evaluations to estimate the error in previous ones. However, that error is precisely what we want to measure, not adjust for it! The results are shown in the following table.

Table 1. Outcome Equities (1ply vs 0ply)

Format	Lose BG	Lose G	Win	Win G	Win BG	**E(ppg)**	SE(ppg)
DMP			0.5148			**0.0296**	0.0008
Cubeless	0.0058	0.1321	0.5108	0.1426	0.0063	**0.0325**	0.0012
Cubeful						**0.0564**	0.0025

2.3. Comparing Money Formats

If we just compared the equities of the better player, cubeful games would appear to be the most skillful. But that could be simply because more points are at stake. Since players will bet less money per point when more points are at stake, we must adjust for the expected value of a game. Assuming players are willing to risk a fixed amount of money per game, it follows that the amount they're willing to bet per point must be inversely proportional to the game's expected value. Therefore, we can use the ratio E/V as a measure of skill.

Table 2. Comparison of Money Formats

Format	E(ppg)	V(ppg)	**E/V**
DMP	0.0296	1.0000	**0.0296**
Cubeless	0.0325	1.2884	**0.0252**
Cubeful	0.0564	2.3856	**0.0236**

These adjusted values tell a different story. Not only are cubeful games the least skillful, but it turns out that DMP games are the most skillful. One possible explanation is that DMP strategy involves more difficult decisions. In contrast, the skill introduced by the cube may be offset by the opportunity for skill lost when cubeful games end with a pass.

3. Match Play

We will compare 3 types of matches:

- cubeful matches with the Crawford rule in effect (cubeful matches),
- cubeless matches with gammons and backgammons (cubeless matches) and
- cubeless matches without backgammons or gammons (DMP matches).

In order to compare the skill in these formats, we again need 2 things for each one:

- the probability of the better player winning an N-point match and
- the expected duration of an N-point match relative to DMP (assuming equal players).

3.1. Defining Skill in Match Play

An obvious way to define the skill S of a match relative to a DMP game is as the ratio of the corresponding expectations of the better player:

$$S = \frac{2P - 1}{2W - 1} \tag{1}$$

where P, W are the probabilities of the better player winning the match or a DMP game respectively. Under this definition, the larger the skill difference of the players is, the smaller the corresponding skill values will be. Another way of defining skill would be using the ELO system. According to the chess ELO formula, the probability of the better player winning a single game is

$$P = \frac{1}{1 + 10^{-\frac{|\Delta R|}{C}}}$$

where R is a player's rating and C is a constant. The only difference in backgammon is that the ELO difference at DMP is multiplied by a factor depending on the match length to adjust for the greater advantage the better player has in longer matches. This factor can be defined as the skill of that particular match length. The ELO formula for the better player in cases where the skill isn't constant can be generalized as follows:

$$P(N, W) = \frac{1}{1 + 10^{-\frac{|\Delta R|}{C} S(N, W)}}$$

where W is the probability of the better player winning a DMP game, N is the length of the match and S is a skill function. If we define DMP games to contain 1 unit of skill, we can solve this formula for S(N,W) to get the generalized skill function:

$$S(1, W) = 1 \iff -\frac{C}{|\Delta R|} = \frac{1}{\log\left(\frac{1}{P(1, W)} - 1\right)} = \frac{1}{\log\left(\frac{1}{W} - 1\right)} \implies$$

$$S(N, W) = \frac{\log\left(\frac{1}{P(N, W)} - 1\right)}{\log\left(\frac{1}{W} - 1\right)} \tag{2}$$

Under that definition, the skill values increase with the absolute value of the ELO difference. Note that, although equations (1) & (2) might look very different, they're actually equivalent in the limit as $W \to 1/2$. Since what we're interested in is the opportunity for skill inherent in a match, that is to say the opportunity for skill it gives to the better player regardless of their skill difference, we can define that limit as the skill of an N-point match. If we now find the Match Winning Chances (MWC) of unequal players that are closely matched, we can plug them into equation (2) and calculate the corresponding skill values at different match lengths.

3.2. Expected Duration of Matches

The expected duration D[M,N] from any away score -M-N of a cubeless match can be calculated recursively using the following formula:

$$D[M,N] = T + (1-G)\frac{D[M-1,N]+D[M,N-1]}{2} + (G-B)\frac{D[M-2,N]+D[M,N-2]}{2} + B\frac{D[M-3,N]+D[M,N-3]}{2} \quad (3)$$

where T is the average duration of a game, G is the gammon rate and B is the backgammon rate at that particular score. Since the gammon rate isn't very score-sensitive, we can use Keith's outcome probabilities of money games for all scores farther than 1-away. For 1-away scores, we can extract a **gammon rate of 29.05%** from Kazaross' XG2 Match Equity Table (MET) at -1-2C [2]. For DMP matches, we simply set **G=B=0**. In order to estimate the average duration, we used the time it takes GNUbg 0ply to perform a rollout in a few special cases: DMP which is our unit of measurement, -1-2C for 1-away scores and cubeless games for the rest.

For cubeful matches, the procedure is a bit more complicated. Not only do we need the average duration of every score, but we also need the probability of every possible outcome. Using a similar procedure, Joseph Heled has also constructed cubeful duration tables [3] which can be converted to show the expected number of (DMP) games remaining. Note that Heled's values are smaller than ours likely because he used unequal players whose matches tend to be shorter. Also note that because our calculations for DMP matches require no assumptions, the results are perfectly accurate.

Table 3. Duration of DMP Matches

Length	1	2	3	4	5
Duration	1	2.5	4.12	5.81	7.54

Table 4. Duration of Cubeless Matches

Length	1	2	3	4	5	6	7
Duration	1	1.82	2.92	4.02	5.16	6.32	7.5

Table 5. Duration of Cubeful Matches

Length	1	2	3	4	5	6	7	8	9	10	11	12	13
Zoidis	**1**	**1**	**1.92**	**2.17**	**2.99**	**3.44**	**4.2**	**4.64**	**5.43**	**6.01**	**6.64**	**7.21**	**7.89**
Heled	1	1	1.92	2.15	2.92	3.27	3.96	4.39	5.04	5.46	6.04	6.54	7.20

3.3. Skill in Match Play

The MWC of the better player from any away score -M-N in a DMP match can be calculated recursively using the following formula:

$$P[M,N]=W\cdot P[M-1,N]+(1-W)\cdot P[M,N-1] \tag{4}$$

where W is the win probability of the better player. This formula can be generalized for cubeless matches but, unlike the expected duration, the results are not accurate enough. Therefore, the MWC of the better player were computed by rolling out every score of the 7-point match **38,880 times**.

For cubeful matches, we already had some data. Joseph Heled had used bot self-play to look into the ELO system [4,5]. Because his results are not even close to what the model predicts, we had to roll out the 13-point cubeful match as well – the resulting METs can be found in Appendix 1. Our results also disagree with the theoretical model but are much closer to what we found for cubeless matches. The problem with the system is the skill function $S(N)=\sqrt{N}$ used in backgammon, which was chosen on theoretical grounds rather than empirical evidence [4]. As for Heled's results, the choice of players is most likely the culprit. For example, using noise to obtain a weaker player leads to nonsensical results (longer matches containing less skill than shorter ones). All these results are shown in the following tables. Note that, like the expected duration, the skill values of DMP matches are perfectly accurate.

Table 6. Skill in DMP Matches

Length	1	2	3	4	5
Skill	1.00	1.50	1.88	2.19	2.46
MWC (%)	51.35	52.02	52.53	52.95	53.32

Table 7. Skill in Cubeless Matches

Length	1	2	3	4	5	6	7
Skill	1.00	1.55	1.88	2.08	2.31	2.49	2.69
MWC (%)	51.35	52.09	52.54	52.80	53.11	53.36	53.63

<p style="text-align:center">Table 8. Skill in Cubeful Matches</p>

Length	1	2	3	4	5	6	7	8	9	10	11	12	13
$\sqrt{Length}$	1.00	1.41	1.73	2.00	2.24	2.45	2.65	2.83	3.00	3.16	3.32	3.46	3.61
Heled	1.00	1.00	1.24	1.26	1.45	1.50	1.63	1.67	1.78	1.83	1.92	1.97	2.05
Zoidis	**1.00**	**1.00**	**1.36**	**1.46**	**1.67**	**1.83**	**2.00**	**2.03**	**2.21**	**2.28**	**2.43**	**2.48**	**2.61**
MWC (%)	51.35	51.35	51.84	51.97	52.26	52.47	52.70	52.74	52.98	53.08	53.28	53.35	53.52

If we compare matches of equivalent length, cubeless matches dominate cubeful matches which in turn dominate DMP matches. We say that one format dominates another if matches of the former type are shorter in duration and contain more skill than matches of the latter type. For example, the 2-point cubeless match lasts 1.82 games and has a skill value of 1.55 whereas the 3-point cubeful match lasts 1.92 games and has a skill value of 1.34 only. These differences can be seen most clearly if we plot the skill values against the expected duration. Note that if we use Heled's values for the expected duration, the gap between cubeless and cubeful shrinks but does not close.

Figure 1.
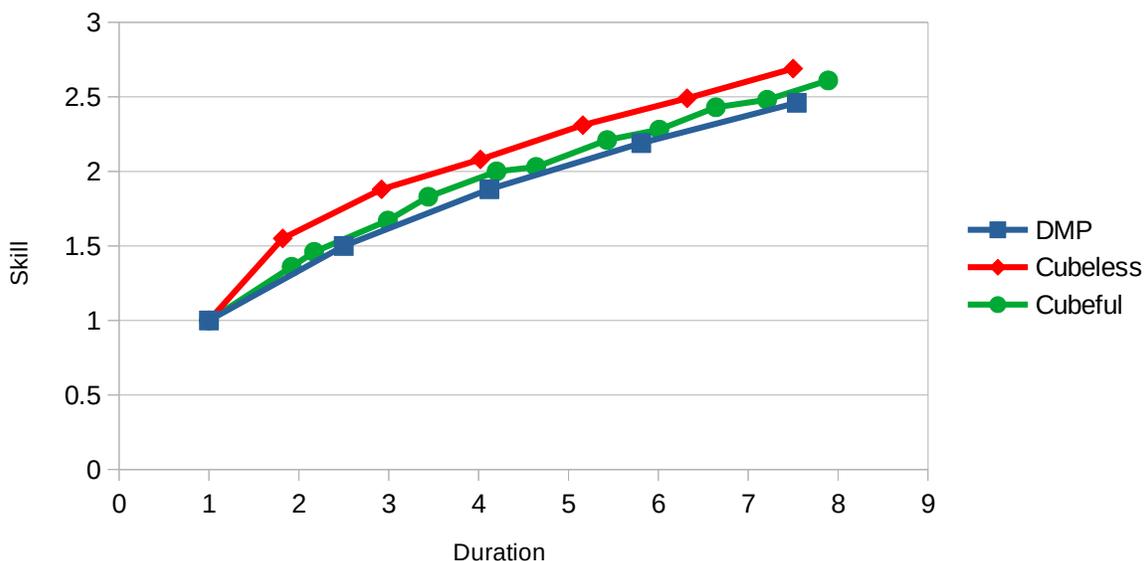
These results shouldn't be surprising. When the stakes of the current game are increased, both players get closer to the end of the match, which means less opportunity for skill. Also, if the match goes beyond the Crawford game, the trailer – who statistically is more likely to be the worse player – has free access to the cube. As for DMP matches, they do much worse than in money games due to the lack of match strategy – the goal is always to win the current game regardless of score.

3.4. The ELO System

It would be useful to have explicit formulas, even approximate, for the skill functions of the various formats to use in the ELO system. For DMP matches in particular, we can derive exact formulas (see Appendices 2 & 3) for both the skill and the expected duration:

$$S(N) = \binom{2N}{N} \frac{2N}{2^{2N}} \tag{5}$$

$$D(N) = 2N - S(N) \tag{6}$$

This last equation reveals a deep connection between skill and expected duration. Assuming luck was evenly distributed, the difference in points at the end of a match represents how much better a player the winner is and can thus be used as a measure of skill. Now simply observe that if the difference in points at the end of an N-point match is S, the duration of the match would be

$$D = N + (N - S) = 2N - S$$

Similar relationships between skill and expected duration exist for all types of backgammon matches. Specifically, the sum of the duration and skill functions seems to always be a linear function of the match length, while the skill values can be fitted with a square root function.

Table 9. Approximate Skill & Duration Formulas

Format	$S(N)$	$D(N)+S(N)$
DMP	$\sqrt{1.27N-0.27}$	$2N$
Cubeless	$\sqrt{1.09N-0.09}$	$1.37N+0.63$
Cubeful	$\sqrt{0.47N+0.53}$	$0.69N+1.31$

Note that the coefficient of N in the skill function approximation for DMP matches agrees with the $\sqrt{4/\pi}$ coefficient of N we get when we apply Stirling's formula to (5). Also note that the square root hypothesis, though inaccurate for cubeful matches, is a surprisingly good fit for cubeless matches. This comeback of the square root is not a coincidence. As we noted above, skill can be represented by the difference in points at the end of a match. In a DMP match that difference corresponds to the distance from the origin in a random walk and since the average distance form the origin is proportional to the square root of the number of steps, it shouldn't be surprising that the square root makes an appearance. After all, this is precisely the reason why it was chosen in the first place.

4. Conclusion

We examined various formats in an effort to find out which one favor the better player most. It turns out that the most skillful forms of play are DMP games and cubeless matches. Now I know that players might object to the idea of parting with their doubling cubes, not to mention playing without gammons, but fortunately for these people the cube will always have its place in chouettes. If however you are one of the many people who complain about luck in tournaments, now you know what the numbers say.

5. References

1. Keith, Tom. "Rollouts of Opening Moves" *Backgammon Galore,* 2006, https://www.bkgm.com/openings/rollouts.html

2. Keith, Tom. "The Kazaross-XG2 Match Equity Table" *Backgammon Galore,* 2013, https://www.bkgm.com/articles/Keith/KazarossXG2MET/index.html

3. "Time to Think - Time to GO (or: what the data says about Mochy's time management)" *YouTube,* uploaded by Joseph Heled, 9 August 2024, https://www.youtube.com/watch?v=z2Jp7GXe5PI

4. "The Backgammon Elo system: theory, practice, data - and why it's (sort of) broken" *YouTube,* uploaded by Joseph Heled, 23 May 2023, https://www.youtube.com/watch?v=Lv75r6xH82Y

5. "PRestigious PR (or: the curious case of Brad Garner)" *YouTube,* uploaded by Joseph Heled, 25 October 2023, https://www.youtube.com/watch?v=pgG8bQgvhI4

A.1. Unequal Skill Equity Tables

CF	PC	-1	-2	-3	-4	-5	-6	-7	-8	-9	-10	-11	-12	-13
PC	**51.35**		52.72	69.92	71.27	82.73	83.66	89.93	90.58	94.14	94.56	96.59	96.86	98.02
-1		**51.35**	69.92	76.43	83.15	85.75	90.15	91.84	94.30	95.28	96.71	97.26	98.08	98.42
-2	49.94	33.39	**51.35**	61.73	69.09	76.49	81.88	85.92	89.11	91.57	93.53	94.98	96.16	97.01
-3	33.39	26.36	41.93	**51.84**	59.60	67.18	73.63	78.52	82.81	86.07	89.00	91.15	93.05	94.48
-4	32.13	19.61	34.60	44.30	**51.97**	59.89	66.60	72.25	76.98	81.13	84.53	87.40	89.78	91.70
-5	19.95	16.92	27.37	36.79	44.40	**52.26**	59.25	65.26	70.56	75.21	79.26	82.74	85.66	88.19
-6	18.96	12.00	21.37	30.43	37.65	45.63	**52.47**	58.86	64.37	69.60	74.15	78.04	81.48	84.43
-7	11.44	10.05	17.03	25.32	31.96	39.60	46.18	**52.70**	58.34	63.76	68.68	73.01	76.82	80.27
-8	11.21	6.94	13.46	20.81	27.07	34.16	40.61	47.05	**52.74**	58.37	63.38	68.05	72.22	75.97
-9	7.19	5.88	10.73	17.21	22.75	29.39	35.41	41.74	47.35	**52.98**	58.12	62.95	67.37	71.52
-10	6.83	4.29	8.42	14.06	19.09	25.07	30.83	36.80	42.31	47.93	**53.08**	58.07	62.60	66.90
-11	4.00	3.62	6.72	11.60	15.99	21.45	26.62	32.36	37.65	43.10	48.29	**53.28**	57.93	62.35
-12	3.96	2.45	5.25	9.34	13.31	18.18	22.95	28.30	33.36	38.68	43.65	48.72	**53.35**	57.92
-13	2.53	2.07	4.18	7.66	11.05	15.39	19.73	24.59	29.46	34.43	39.37	44.27	48.94	**53.52**

CL	-1	-2	-3	-4	-5	-6	-7
-1	**51.35**	69.92	82.73	89.93	94.14	96.59	98.02
-2	33.39	**52.09**	67.66	78.39	85.81	90.81	94.11
-3	19.95	36.58	**52.54**	65.33	75.32	82.75	88.14
-4	11.44	25.07	39.62	**52.80**	64.07	73.28	80.53
-5	7.19	17.10	29.19	41.54	**53.11**	63.25	71.84
-6	4.00	11.36	21.00	31.91	42.96	**53.36**	62.75
-7	2.53	7.63	14.99	24.12	34.06	44.05	**53.63**

A.2. The DMP Skill Formula

$$S(N) = \lim_{W \to 1/2} S(N, W) = \lim_{W \to 1/2} \frac{\log\left(\dfrac{1}{P(N, W)} - 1\right)}{\log\left(\dfrac{1}{W} - 1\right)}$$

Since an N-point cubeless DMP match is equivalent to a best of 2N-1 match, we can use the binomial distribution to calculate the probability of the better player winning:

$$P(N, W) = \sum_{K=N}^{2N-1} \binom{2N-1}{K} W^K (1-W)^{2N-1-K}$$

Using L' Hopital's rule we get

$$S(N) = \lim_{W \to 1/2} \frac{(1-W) \cdot W}{[1-P(N,W)] \cdot P(N,W)} \sum_{K=N}^{2N-1} \binom{2N-1}{K} \cdot W^{K-1} \cdot (1-W)^{2N-2-K} \cdot [K - (2N-1) \cdot W]$$

Since $P(N, 1/2) = 1/2$ we have

$$\sum_{K=N}^{2N-1} \binom{2N-1}{K} = P(N, 1/2) \cdot 2^{2N-1} = 2^{2N-2} \text{ and our limit becomes}$$

$$S(N) = 2^{3-2N} \cdot \sum_{K=N}^{2N-1} \binom{2N-1}{K} \cdot K - \frac{2N-1}{2^{2N-2}} \sum_{K=N}^{2N-1} \binom{2N-1}{K} \iff$$

$$S(N) = (2N-1) \cdot 2^{3-2N} \cdot \sum_{K=N}^{2N-1} \binom{2N-2}{K-1} - (2N-1)$$

$$\text{Let } A = \sum_{K=N}^{2N-1} \binom{2N-2}{K-1} = \sum_{K=N-1}^{2N-2} \binom{2N-2}{K} = \sum_{K=0}^{N-1} \binom{2N-2}{K} \Rightarrow$$

$$2A = \sum_{K=0}^{N-1} \binom{2N-2}{K} + \sum_{K=N-1}^{2N-2} \binom{2N-2}{K} = \binom{2N-2}{N-1} + \sum_{K=0}^{2N-2} \binom{2N-2}{K} = \binom{2N}{N} \frac{N}{2 \cdot (2N-1)} + 2^{2N-2} \Rightarrow$$

$$S(N) = \binom{2N}{N} \frac{2N}{2^{2N}} \quad \blacksquare$$

A.3. The DMP Duration Formula

Since an N-point cubeless DMP match will end after a minimum of N games and a maximum of 2N-1 games, the expected duration can be expressed as:

$$D(N)=\sum_{K=N}^{2N-1} K \cdot P(K)$$

where P(K) is the probability of the match lasting exactly K games. In order for the match to end after K games, the winner of game K must win N-1 of the first K-1 games. There are C(K-1,N-1) combinations in which this happens. Assuming both players are equally likely to win a single game, the probability of each combination is 2^{1-K} and thus we have:

$$D(N)=\sum_{K=N}^{2N-1} K \cdot \binom{K-1}{N-1} \cdot 2^{1-K} = N \sum_{K=N}^{2N-1} \binom{K}{N} \cdot 2^{1-K} \Leftrightarrow$$

$$D(N)=N \sum_{K=0}^{N-1} \binom{K+N}{N} 2^{1-K-N} = N \cdot 2^{1-2N} \sum_{K=0}^{N-1} \binom{N+K}{N} \cdot 2^{N-K} \Leftrightarrow$$

$$D(N)=N \cdot 2^{1-2N} \{ \sum_{K=0}^{N} \binom{N+K}{N} \cdot 2^{N-K} - \binom{2N}{N} \}$$

To simplify our calculations let's consider a best of 2N+1 match, equivalent to an (N+1)-point match. Let K be the score of the loser of the match. The game that decides the match is preceded by exactly N games won by the winner and K games won by the loser. These N+K games can occur in any order. Now imagine that the maximum of 2N+1 games are played even if the winner's already decided. Thus, there remain N-K games which can go either way. By symmetry, assuming both players are equally likely to win, the total number of sequences of 2N+1 games is equal to twice the total number of sequences that decide the outcome in the (N+K+1)-st game. Therefore, we have:

$$2 \sum_{K=0}^{N} \binom{N+K}{K} \cdot 2^{N-K} = 2^{2N+1} \Leftrightarrow \sum_{K=0}^{N} \binom{N+K}{N} \cdot 2^{N-K} = 2^{2N} \Rightarrow$$

$$D(N)=2N \cdot \{1 - \binom{2N}{N} \cdot 2^{-2N}\} \Leftrightarrow$$

$$D(N)=2N-S(N) \quad \blacksquare$$