BONUS! Maximizing Surprise

Zhihuan Huang\textsuperscript{1}, Yuqing Kong\textsuperscript{*1}, Tracy Xiao Liu\textsuperscript{†2}, Grant Schoenebeck\textsuperscript{‡3}, and Shengwei Xu\textsuperscript{1}

\textsuperscript{1}Department of Computer Science, Peking University
\textsuperscript{1}Center on Frontiers of Computing Studies, Peking University
\textsuperscript{2}School of Economics and Management, Tsinghua University
\textsuperscript{3}School of Information, University of Michigan
\texttt{\{zhihuan.huang, shengwei.xu, yuqing.kong\}@pku.edu.cn}
\texttt{liuxiao@sem.tsinghua.edu.cn}
\texttt{schoeneb@umich.edu}

Abstract

Multi-round competitions often double or triple the points awarded in the final round, calling it a bonus, to maximize spectators’ excitement. In a two-player competition with \(n\) rounds, we aim to derive the optimal bonus size to maximize the audience’s overall expected surprise (as defined in [7]). We model the audience’s prior belief over the two players’ ability levels as a beta distribution. Using a novel analysis that clarifies and simplifies the computation, we find that the optimal bonus depends greatly upon the prior belief and obtain solutions of various forms for both the case of a finite number of rounds and the asymptotic case. In an interesting special case, we show that the optimal bonus approximately and asymptotically equals to the “expected lead”, the number of points the weaker player will need to come back in expectation. Moreover, we observe that priors with a higher skewness lead to a higher optimal bonus size, and in the symmetric case, priors with a higher uncertainty also lead to a higher optimal bonus size. This matches our intuition since a highly asymmetric prior leads to a high “expected lead”, and a highly uncertain symmetric prior often leads to a lopsided game, which again benefits from a larger bonus.

\textsuperscript{*}Supported by National Natural Science Foundation of China award number 62002001
\textsuperscript{†}Supported by the National Key Research and Development Program of China award number 2018YFB1004503
\textsuperscript{‡}Supported by (United States) National Science Foundation award number 2007256
1 Introduction

People love watching competitions. Who will snatch victory? Who will be vanquished? Thus, the business of sports, e-sports, and live streaming platforms create a huge amount of revenue.

When people watch a game, their belief for who would win in the end changes over the duration of the game. Intuitively, the surprise, measures how much audience’s beliefs change over time [7]. An important question for both theorists and practitioners is how to design winner selection schemes that maximize the amount of surprise in a competition and, consequently, improve the entertainment utility of a competition and increase its potential revenue.

In practice, many games use point systems to determine the winner. For some games, the point value remains static throughout the game, such as soccer, cricket, and tennis. However, in other games, the point value in the final round is higher than in others. Mind King, a very popular two-player quiz game on WeChat has 5 rounds. For each round, each player receives points depending on the correctness and speed of their answer. The final round doubles the points. Since 2015, IndyCar racing has doubled the points for the final race of the season. The Diamond League, a track and field league, from 2010-2016, determined the Diamond Race winner by a point system over a season of 7 meets, and the points for the final tournament are doubled. Another very popular example is the quidditch matches in Harry Potter. The game concludes when the golden snitch, which is worth 15 times a normal goal, is captured by one team. While this point structure makes sense for the plot of the books (the title character often catches the snitch), most of the action is ancillary to the games’ outcome. Tellingly, “muggle” quidditch is now a real-life game, however, the golden snitch is only worth 3 times as much as a normal goal.

Key Question In a simple model, we study the effect of the final round’s point value on the overall surprise and what point value maximizes the surprise. In our setting, there are two teams and a fixed number of rounds. The winner in each of the first \(n - 1\) rounds is awarded 1 point and the winner in the final round can possibly receive more points, e.g., double or triple the points earned in previous rounds. The team which accumulates the most points wins.

---

1FIFA generated more than 4.6 billion USD revenue in 2018, mainly on 2018 FIFA World Cup, according to https://www.investopedia.com/articles/investing/070915/how-does-fifa-make-money.asp

2Douyu, one of the major e-sports live streaming platform in China generated about 6.6 billion CNY of revenue (around 1 billion USD) in 2019, according to https://www.statista.com/statistics/1222790/china-douyu-live-streaming-revenue/

3The worldwide video streaming market size reached 50.11 billion USD in 2020, according to https://www.grandviewresearch.com/industry-analysis/video-streaming-market

4over 1 million daily users according to an author interview with Tencent staff in July 2021.

5According to http://www.champcarstats.com/points.htm

6According to Wikipedia: https://en.wikipedia.org/wiki/Diamond_League#Scoring_system

7The scoring system was substantially changed in 2017 to, in particular, include a final restricted to the top-scoring athletes.

8Harry Potter is a very popular fantasy novels series that have sold out more than 500 million copies according to https://www.wizardingworld.com/news/500-million-harry-potter-books-have-now-been-sold-worldwide
Types of Prior Beliefs  To measure surprise, we need to model audience belief. The audience’s belief of who will win (and thus the surprise) depends on the audience’s prior belief about the two contestants’ chances of winning in each round. Note that this belief may be updated as the match progresses. Intuitively, in a game where audiences believe that the two players’ ability is highly asymmetric, e.g., a strong vs. a weak player, we should set up a high bonus score, otherwise, the weaker player will likely be eliminated before the game’s conclusion. In contrast, if the two players are perfectly evenly matched and the total number of rounds is odd, we will show the optimal bonus value is 1, the same point value as in the previous rounds. Below, we consider three special cases for prior beliefs and then introduce the general case.

The first case is that the audience has a fixed and unchanging belief about the chance that each competitor wins for each round. We call this the certain case. The size of the optimal bonus depends on the belief in the difference between two contestants’ ability levels.

The second case is where the audience has no prior knowledge about the two players’ abilities. We call this the uniform case. Specifically, we model the audience’s prior over the two players as a uniform distribution and derive the optimal bonus size accordingly. A slight generalization of the uniform case is the symmetric case when the audience has a certain amount of prior knowledge but no prior knowledge favoring one contestant over another. In this case, we model the audience’s prior as a symmetric Beta distribution. In the general case, we model the prior as a general Beta distribution, \( \text{Be}(\alpha, \beta) \), a family of continuous probability distributions on the interval \([0, 1]\) parameterized by two positive shape parameters, \( \alpha, \beta \geq 1 \). The figures above Table 1 illustrate the above cases.

Techniques  Our analysis is built on two insights. First, for Beta prior beliefs, we show that for \( i \leq n - 1 \), the ratio of the expected surprise in round \( i \) and \( i - 1 \) is a constant that only depends on \( i \) and the prior. This allows us to reduce the entire analysis to the trade-off between the final round’s surprise amount and the penultimate round’s surprise amount. Second, we find that the final round’s surprise amount increases when the bonus increases while the penultimate round’s surprise consists of two parts where one part increases with the bonus size and the other part decreases with the bonus size. Optimally trading off these terms yields our results. Third, we show how to optimize this trade-off.

1.1 Results

Table 1 shows the optimal bonus size in each case of the prior—symmetric, certain, and general—for both any finite number of rounds and the asymptotic value as the number of rounds increases. Note that we have closed-form formulae for the symmetric case and the asymptotic certain case.

We obtain the following interesting insights:

Insight 1. In the certain case, more uneven match-ups lead to a larger optimal bonus  When the match is more uneven, i.e., when the prior has a higher skewness \( \frac{\alpha - \beta}{\alpha + \beta} \),

---

9It is worthwhile to note that this is different from the case where the prior belief is that both players have an equal chance to win. The reason is that no belief updating exists in the prior case while belief updating exists in the later one.

10When Alice and Bob have played before and Alice wins \( n_A \) rounds and Bob wins \( n_B \), we can use \( \text{Be}(n_A + 1, n_B + 1) \) to model the prior belief. In general, we allow \( \alpha, \beta \) to be non-integers.
Symmetric  
\[ \alpha = \beta \]  
Certain  
\[ \alpha = \lambda p, \beta = \lambda(1 - p), \lambda \to \infty \]  
General  
\[ \alpha = \beta \]

| Finite | Asymptotic |
|--------|------------|
| \( \text{RD}(\frac{n - 1}{2\alpha H - \frac{n - 1}{n + 2\alpha - 1}}) \) | \( n \text{ rd}(\frac{1}{2\alpha H - 1} \approx \frac{n}{2\alpha} \) |
| \( \text{RD}(\text{Solution of } F(x) = 0) \) | \( n \frac{\alpha - \beta}{\alpha + \beta} \) |
| \( O(n) \) algorithm | \( n \text{ rd}(\frac{n(\alpha - \beta)H + 1}{(\alpha + \beta)H - 1} \approx \frac{n}{2\alpha} \) |

**Table 1: Optimal bonus**

Without loss of generality, we assume \( \alpha \geq \beta \) which implies that \( p \geq \frac{1}{2} \).

\[
\text{RD}(x) \approx \text{the nearest integer to } x \text{ that has the same parity as the number of rounds } n\]

\[
H = H_{\alpha+\beta}(n-1) := \sum_{i=1}^{n-1} \frac{1}{i + \alpha + \beta - 1},
\]

\[
F(x) := (2np - n - (x-1))p^{x-1} + (n - 2np - (x-1))(1-p)^{x-1}, x \in [1, n+1].
\]

\[
G(\mu) := (1 + \mu)^{\alpha - \beta} \left( \frac{(\alpha - \beta)H + 1}{(\alpha + \beta)H - 1} - \mu \right) + (1 - \mu)^{\alpha - \beta} \left( \frac{(-\alpha + \beta)H + 1}{(\alpha + \beta)H - 1} - \mu \right), \mu \in (0, 1)
\]

\( ^a \)When there is a tie, we pick the smaller one.

\( ^b \)When \( p = \frac{1}{2} \) or \( n < \frac{1}{(2 - p) \ln(\frac{1}{2p})} \), \( F(x) \) only has a trivial solution \( x = 1 \) and the optimal bonus is \( \text{RD}(1) \), otherwise, the optimal bonus is \( \text{RD}(\tilde{x}) \) where \( \tilde{x} \) is the unique non-trivial solution.

The optimal bonus is larger. Interestingly, in the certain case, we find that the optimal bonus is around \( \frac{\alpha - \beta}{\alpha + \beta} n \) which is the expected amount of points the weaker player needs to come back, which we call the “expected lead”. For example, we let Marvel superheroes compete in multi-rounds. When Thor fights Hawkeye, there should be a large bonus than when Thor fights Iron Man.

**Insight 2. In the symmetric case, more uncertainty leads to a higher optimal bonus.** In the symmetric case when the prior has more uncertainty, i.e., when \( \frac{1}{\alpha + \beta} = \frac{1}{2\alpha} \) is larger, the optimal bonus is larger. In particular, the uniform case has a higher bonus than the certain case. For example, if Thor fights Superman, no one really knows the relative abilities of the heroes, since they come from different worlds. There should be a larger bonus than if Thor fights Iron Man. Since Thor has fought Iron Man before and it is known that they will be a good match-up. Note that in general, the optimal bonus size is not monotone in the uncertainty while holding the skewness fixed.

Figure 1 shows the optimal bonus’s for the asymptotic case. For small \( \alpha - \beta \), it is similar to the symmetric case (the y-axis); and, analogously, for large \( \alpha - \beta \), it is similar to the certain case (the x-axis). Thus, we can use the special cases’ results to approximate the general case.
Figure 1: **Optimal bonus in asymptotic case:** for all $\alpha, \beta \geq 1$, we use $\frac{\alpha - \beta}{\alpha + \beta}$ to measure the skewness of the prior and $\frac{1}{\alpha + \beta}$ to measure the uncertainty of the prior. The figure shows that when $n$ is sufficiently large (we use $n = 10000$ here), the relationship between the optimal bonus size and skewness/uncertainty.

1.2 Related Work

Ely et al. [7] provide a clear definition of surprise amount. Starting from Ely et al. [7], a growing literature examines the relationship between the surprise and the perceived quality in different games and presents empirical support that audiences’ perceived quality of a game is partly determined by the amount of surprise generated in this process, such as tennis [2], soccer [5,11], and rugby [14]. Ely et al. [7] additionally study what number of rounds maximizes surprise in our certain case. In contrast, we study how to derive the optimal number of points for the final round and consider a more general prior model.

In addition to the surprise-related work, there has been a growing number of theoretical and empirical results studying how game rules affect different properties of the game, primarily fairness. For example, Brams and Ismail [3] study how to make penalty shootouts rule fairer. Braverman et al. [4] study how to make a popular party game, mafia, fair by tuning the number of different characters in the game. Percy [13] shows that the change of badminton scoring system in 2006, in an attempt to make the game faster paced, did not adversely affect the fairness.

One conception of fairness corresponds to the “better” team winning. However, in some cases, like the certain case, this makes for a very dull game as there will be no surprise possible. Thus, the goal of optimizing surprise and making the “better” team consistently win are sometimes in tension.

In addition to fairness, Percy [13] also studies the influence of changes in the badminton scoring system implemented in 2006 on the entertainment value which is related to the number of rallies. However, in contrast to our work, Percy [13] focuses on comparing two
scoring rules rather than optimization and does not formally model the entertainment value. Kovacs [9] studies the effect of changes in the volleyball scoring system and empirically shows that the change may make the length of the matches more predictable.

Results from Mossel et al. [12] based on Fourier analysis find that if one team wins each round with probability \( p > 1/2 \) then they win the match with probability at least \( p \). Moreover, the only way for the winning probability to be exactly \( p \) is for the match to be decided by the outcome of one round. (They were studying opinion aggregating on social networks and this result applies to all deterministic symmetric and monotone Boolean functions with i.i.d. inputs.) This directly relates to the present work because, ideally, each team would start with a prior probability of winning close to 1/2. This result says that the dictator function (which can be enacted in our setting with a bonus of size \( n \)) yields a prior as close to 1/2 as possible. Our results in the certain case can be seen as trading off the two goals of having a prior probability close to 1/2 but also revealing information more slowly over time than a dictator function.

## 2 Problem Statement

We consider an \( n \) round competition between two players, Alice and Bob. In round \( i \), a task is assigned to the players and the winner receives points. After the \( n \)th round, the player with the higher accumulated score wins the competition. In our setting, we assume that each of the first \( n-1 \) rounds is worth one point. However, the point value \( x \) for the final round might be different, and we call \( x \) the *bonus*. Setting the final round bonus is a special, but interesting case to study. Without loss of generality, we consider \( x \) to be an integer where \( 0 \leq x \leq n \).

To ensure there is no tie, we additionally require \( x \) has the same parity as \( n \). Let \( \mathcal{X}(n) \) represent our considered range for \( x \), that is

\[
\mathcal{X}(n) = \begin{cases} 
\{1, 3, 5, \ldots, n\}, & \text{when } n \text{ is odd} \\
\{0, 2, 4, \ldots, n\}, & \text{when } n \text{ is even}
\end{cases}
\]

### 2.1 Optimization Goal

To introduce our optimization goal formally, we first introduce the concept of belief curve.

**Belief Curve** The audience’s belief curve is a sequence of random variables

\[ B := (B_0, B_1, \ldots, B_n) \]

where \( B_i, i \in [n] \) is the audience’s belief for the probability that Alice wins the whole competition after round \( i \). \( B_0 \) is the initial belief. \( B_n \) is either zero or one since the final outcome must be revealed in the end.

We define \( O \in \{0, 1\} \) as the outcome of the whole competition, that is

\[
O = \begin{cases} 
0 & \text{Alice loses the whole competition} \\
1 & \text{Alice wins the whole competition}
\end{cases}
\]

\[ ^1 \text{Mathematically, } x > n \text{ is equivalent to } x = n \text{ where only the final round matters. Any non-integer } x \text{ is equivalent to an integer. For example, when } n = 4, x = 1.3 \text{ is equivalent to } x = 2. \]
then we define random variables

\[ H_i = \begin{cases} - & \text{Alice loses } i\text{-th round} \\ + & \text{Alice wins } i\text{-th round} \end{cases} \]

\[ \mathcal{H}^{(i)} = (H_1, H_2, \ldots, H_i) \]

i.e., \( \mathcal{H}^{(i)} \) consists of the history of the first \( i \) rounds, we call \( \mathcal{H}^{(i)} \) an \( i \)-history.

Next, we define random variables

\[ B_i = \Pr[O = 1 | \mathcal{H}^{(i)}] \]

\[ \mathcal{B}^{(i)} = (B_0, B_1, \ldots, B_i) \]

i.e., \( B_i \) is the conditional probability that Alice wins the whole competition. Additionally, we use \( \mathcal{B} \) as a shorthand for \( \mathcal{B}^{(n)} \). Note that \( \mathcal{B} \) is a doob martingale [6] thus \( \forall i, \mathbb{E}[B_{i+1} | \mathcal{H}^{(i)}] = B_i \).

**Definition 2.1 (Surprise).** [7] Given the belief curve \( \mathcal{B} \), we define \( \Delta_i^{\mathcal{B}} := |B_i - B_{i-1}| \) as the amount of surprise generated by round \( i \). We define the overall surprise for a given belief curve to be

\[ \Delta_{\mathcal{B}} := \sum_i \Delta_i^{\mathcal{B}}. \]

**Maximizing the Overall Expected Surprise**

We aim to compute the bonus size \( x \) which, in expectation, maximizes the audience’s overall surprise. That is, we aim to find \( x \) to maximize the overall surprise

\[ \arg \max_{x \in \mathcal{X}^{(n)}} \mathbb{E}[\Delta_{\mathcal{B}}(x)] \]

where \( \Delta_{\mathcal{B}}(x) \) is the overall surprise when the bonus round’s point value is equal to \( x \).

**2.2 Model of Prior Belief**

We introduce a natural model for the audience’s prior. We will maximize the overall expected surprise in this model.

We assume that each player’s winning probability across rounds is constant, i.e., Alice wins with the same probability \( p \) in each round. Moreover, we assume the outcomes of these tasks are independent.

**Prior over \( p \)**

We use Beta distribution \( \text{Be}(\alpha, \beta) \) to model the audience’s prior about Alice’s winning probability \( p \) in each round. The family of Beta distributions is sufficiently rich to cover a variety of important scenarios including the uniform case (\( \alpha = \beta = 1 \)), the symmetric case (\( \alpha = \beta \)) and the certain cases (\( \alpha = \lambda p, \beta = \lambda (1 - p), \lambda \rightarrow \infty \)). A key property of the Beta distribution is that, if \( p \) is drawn from a Beta distribution and then we see the outcome of a coin which lands heads with probability \( p \), the posterior of \( p \) after observing a coin flip is still a Beta distribution.

**Claim 2.2 (Beta’s posterior is Beta).** If the prior \( p \) follows \( \text{Be}(\alpha, \beta) \), then conditioning on Alice winning the first round, the posterior distribution over \( p \) is \( \text{Be}(\alpha + 1, \beta) \), and conditioning on Alice losing the first round, the posterior distribution over \( p \) is \( \text{Be}(\alpha, \beta + 1) \).
Let $p|\mathcal{H}^{(i)}$ be a random variable which follows the posterior distribution of $p$ conditioning on i-history $\mathcal{H}^{(i)}$. For all $i \leq n - 1$, we define an induced random variable $S_i := \text{COUNT}(\mathcal{H}^{(i)})$ as the number of rounds Alice wins in the first $i$ rounds, called the state after $i$ rounds. $p|S_i$ is a random variable which follows the posterior distribution of $p$ conditioning on state $S_i$. The property of Beta distribution (Claim 2.2) directly induces the following claim.

**Claim 2.3 (Order Independence).** For all $i \leq n - 1$, for all $h \in \{+, -\}^i$, $p(\mathcal{H}^{(i)} = h)$ follows distribution $\text{Be}(\alpha + \text{COUNT}(h), \beta + i - \text{COUNT}(h))$.

This immediately implies the following corollary.

**Corollary 2.4 (State Dependence).** For all $i \leq n - 1$, for all $h \in \{+, -\}^i$, $p(\mathcal{H}^{(i)} = h)$ follows the same distribution as $p(S_i = \text{COUNT}(h))$ and $\Pr[O = 1|\mathcal{H}^{(i)} = h] = \Pr[O = 1|S_i = \text{COUNT}(h)]$.

That is, the posterior distribution of $p$ or $O$ only depends on the state, i.e., the number of rounds Alice wins and the order does not matter. For example, history $++-$ and history $-++$ induce the same posterior.

The above properties make our prior model tractable. In this model, the optimal bonus size $x^*$ depends on $\alpha, \beta, n$ thus is denoted by $x^*(\alpha, \beta, n)$.

### 2.3 Method Overview

**Technical Challenge** The key technical challenge is that we do not have a clean format for the belief value across all rounds (especially early rounds). For example, for the asymmetric case, it is even difficult to represent the initial belief $B_0$ for different $x$. A naive way to compute all belief values is using backward induction. We can then enumerate all possible $x$ to find the optimal one. However, this method has $O(n^3)$ time complexity.
To overcome this challenge, we utilize the properties of the Beta distributions. First, we show that we only need to analyze the belief values in the final two rounds and choosing \( x \) becomes a trade-off between the final and penultimate rounds\(^{12}\). Second, we study a few important special cases (asymptotic, symmetric, certain) which can further simplify the final two rounds’ analysis significantly. Third, instead of actually computing \( E[\Delta_B(x)] \), we only analyze how \( E[\Delta_B(x)] \) changes with \( x \).

**Method Overview** Our method has three steps.

---

**Step 1 (Main Technical Lemma)** We show that fixing \( n, \alpha, \beta, \forall \ x \), there exists a constant \( C \) such that

\[
E[\Delta_B(x)] = E[\Delta_{B}^{n-1}(x)] * C + E[\Delta_B^n(x)]
\]

Thus, the choice of \( x \) only depends on the trade-off of the penultimate and the final round’s surprise. This significantly simplifies our analysis since we have a close form expression for the belief in the final two rounds.

\(^{12}\)When \( x \) increases, the final round generates more surprise while the penultimate round generates less surprise.
Step 2 (Final & Penultimate Rounds) We define $\Delta^i_{H(i-1)}$ as the expected amount of surprise generated by round $i$ given the history $H(i-1)$. $\Delta^i_{S(i-1)}$ is the amount of surprise generated by round $i$ given that Alice wins $S_{i-1}$ rounds in the first $i - 1$ rounds. Corollary 2.4 shows that $\Delta^i_{H(i-1)}$ only depends on the state of the history, that is, for any history $h \in \{+, -\}^{i-1}$, $\Delta^i_{S_{i-1}=\text{COUNT}(h)} = \Delta^i_{H(i-1)=h}$.

The final round: In the final round, we notice when the difference between Alice and Bob is strictly greater than $x$, the outcome of the whole competition does not change regardless of who wins the final round, and no surprise is generated in the final round. Formally, we define $L_{n-1} := \frac{n-x}{2}$ and $U_{n-1} := \frac{n+x-2}{2}$. Only states $S_{n-1} \in [L_{n-1}, U_{n-1}]$ generate surprise. Thus, the total surprise generated in round $n$ is

$$E[\Delta^n_B(x)] = \sum_{j=L_{n-1}}^{U_{n-1}} \Pr[S_{n-1} = j] \ast \Delta^n_{S_{n-1} = j}$$

(1)

In any state $S_{n-1} = j \in [L_{n-1}, U_{n-1}]$, whoever wins the final round wins the whole competition and the analysis for all $\Delta^n_{S_{n-1} = j}$ is identical.

The penultimate round: In the penultimate round, similarly, we define $L_{n-2} := \frac{n-x}{2}$ and $U_{n-2} := \frac{n+x-2}{2}$. Similarly, only states $S_{n-2} \in [L_{n-2}, U_{n-2}]$ generate surprise. The states in $(L_{n-2}, U_{n-2})$ are similar. However, unlike the final round, here the states at the endpoints require different analysis. Therefore, we divide the analysis into three parts

$$E[\Delta^{n-1}_{B}(x)] = \sum_{j=L_{n-2}+1}^{U_{n-2}} \Pr[S_{n-2} = j] \ast \Delta^{n-1}_{S_{n-2} = j}$$

(2)

Step 3 (Local Maximum): To calculate the optimal bonus $x$, we need to analyze how $E[\Delta_B(x)]$ changes with $x$.

Finite Case Since we require the bonus $x$ to be an integer with the same parity as $n$, we calculate the change of the function $E[\Delta_B(x)]$ when $x$ is increased/decreased by a step size 2. We find that there is only one local (and thus global) maximum.

Asymptotic Case We calculate the derivative of $E[\Delta_B(x)]$ with respect to $x$, and find that it only has one zero solution which is a local (and also global) maximum.

3 Main Technical Lemma

In this section, we introduce our main technical lemma: regardless of the final round’s point value $x$, the expected surprise in the first $n - 1$ rounds has a fixed relative ratio. Thus, the
overall surprise can be rewritten as a linear combination of the final and the penultimate rounds’ expected surprise, where the coefficients are independent of $x$. This simplifies our analysis significantly since the choice of $x$ only depends on the trade-off between the final and the penultimate rounds’ expected surprise.

**Lemma 3.1** (Main Technical Lemma). When the prior is $\mathcal{B}e(\alpha, \beta)$, the ratio of the surprise of round $i$ and round $i + 1 (i + 1 < n)$ is independent of the final round’s point value $x$:

$$\frac{\mathbb{E}[\Delta^i_B]}{\mathbb{E}[\Delta^{i+1}_B]} = \frac{i + \alpha + \beta}{i + \alpha + \beta - 1},$$

thus the overall surprise is a linear combination of the final and the penultimate round’s surprise,

$$\mathbb{E}[\Delta_B] = \sum_{i=1}^{n} \mathbb{E}[\Delta^i_B] = \mathbb{E}[\Delta^{n-1}_B] \ast (n + \alpha + \beta - 2) \ast \mathbb{H}_{\alpha+\beta}(n-1) + \mathbb{E}[\Delta^n_B]$$

where $\mathbb{H}_{\alpha+\beta}(n-1) := \sum_{i=1}^{n-1} \frac{1}{i + \alpha + \beta - 1}$. We use $\mathbb{H}$ as shorthand for $\mathbb{H}_{\alpha+\beta}(n-1)$.

**Proof Sketch** To prove this lemma, we first introduce Claim 3.2 which gives a simple format of the expected surprise generated by a single round initialized from any history $h$. We then apply the claim to analyze two consecutive rounds initialized from any history and show that the expected surprise produced by these two rounds has a fixed relative ratio which only depends on the round number, $\alpha$, and $\beta$. We then extend the results to the expectation over all possible histories.

To simplify the notation in the proof, we introduce shorthand notations for expectation of $p$, the belief values, and the difference between belief values here:

$$\begin{align*}
q & := \mathbb{E}[p|\mathcal{H}^{(i-1)} = h] \\
q^+ & := \mathbb{E}[p|\mathcal{H}^{(i)} = h^+] \\
q^- & := \mathbb{E}[p|\mathcal{H}^{(i)} = h^-]
\end{align*}$$

$$\begin{align*}
b & := \Pr[O = 1|\mathcal{H}^{(i-1)} = h] \\
b^+ & := \Pr[O = 1|\mathcal{H}^{(i)} = h^+] \\
b^- & := \Pr[O = 1|\mathcal{H}^{(i)} = h^-]
\end{align*}$$

$$\begin{align*}
d & := b^+ - b^- \\
d^+ & := b^{++} - b^{+-} \\
d^- & := b^{+-} - b^{--}
\end{align*}$$

where $b^{++}, b^{+-}, b^{+-}, b^{--}$ are defined analogously. These notations are also illustrated in Figure 4 and Figure 5.

**Claim 3.2.** Given any history $\mathcal{H}^{(i-1)} = h$, we have

$$\Delta^{i}_{\mathcal{H}^{(i-1)} = h} = d \ast 2q(1 - q).$$

**Proof of Claim 3.2** The expected amount of surprise generated in round $i$ is

$$\Delta^{i}_{\mathcal{H}^{(i-1)} = h} = q \ast (b^+ - b) + (1 - q) \ast (b - b^-)$$

(3)

Since $\mathcal{B}$ is a martingale, we have

$$\begin{align*}
B_i & = \mathbb{E}[B_{i+1}|\mathcal{H}^{(i-1)} = h] \\
b & = q \ast b^+ + (1 - q) \ast b^-
\end{align*}$$

(4)

By substituting $b$ from (4) into (3) and simplifying the equation, we have

$$\Delta^{i}_{\mathcal{H}^{(i-1)} = h} = d \ast 2q(1 - q)$$

$\square$
Proof of Lemma 3.1. Figure 5 shows two consecutive rounds starting from history $H(i-1) = h$. We show that the ratio between these two rounds’ surprise is fixed.

By Claim 3.2, we have

$$\Delta_{h(i-1) = h}^i = d * 2q(1 - q)$$

and the expected amount of surprise generated in round $i + 1$ is

$$E[\Delta_{H(i) | H(i-1) = h}^{i+1}] = q * \Delta_{H(i) = h+}^{i+1} + (1 - q) * \Delta_{H(i) = h-}^{i+1} = q * d^+ * 2q^+(1 - q^+) + (1 - q) * d^- * 2q^-(1 - q^-)$$

Properties of Beta distribution imply that the belief given any history only depends on the state of the history (Claim 2.3), thus $b^{++} = b^{--}$. Then we have $d = d^+ * q^+ + d^- * (1 - q^-)$ (see Figure 5) and
Thus, we have

\[
\frac{\Delta^i_{H(i-1) = h}}{E[\Delta^{i+1}_{H(i)}|H(i-1) = h]} = \frac{d \cdot 2q(1-q)}{q \cdot d^+ 2q^+(1-q^+) + (1-q) \cdot d^- 2q^-(1-q^-)}
\]

\[
= \frac{(d^+ q^+ + d^- (1-q^-)) \cdot 2q(1-q)}{q \cdot d^+ 2q^+(1-q^+) + (1-q) \cdot d^- 2q^-(1-q^-)}
\]

\[
= \frac{d^+ q^+ q \cdot (1-q) + d^- (1-q^-) (1-q) \cdot q}{d^+ q^+ q \cdot (1-q^+) + d^- (1-q^-) (1-q) \cdot q^-}
\]

We observe that if \(q(1-q^+) = \Pr[h+|-] = \Pr[h-+] = (1-q)q^-\), \(\frac{1-q}{q^-} = \frac{q}{q^-}\), and the ratio becomes \(\frac{q}{q^-}\).

The posterior of Beta distribution is still a Beta distribution. We denote the distribution over \(p|H^{(i-1)} = h\) by \(Be(\alpha', \beta')\). Then we have

\[
\begin{cases}
  q = \frac{\alpha'}{\alpha' + \beta'}, \\
  q^- = \frac{\alpha}{\alpha' + \beta' + 1}, \\
  q^+ = \frac{\alpha' + 1}{\alpha' + \beta' + 1}
\end{cases}
\]

Thus, we have \(\Pr[h+|-] = \Pr[h-+]\), and the ratio becomes \(\frac{q}{q^-} = \frac{\alpha' + \beta' + 1}{\alpha' + \beta'}\), i.e.,

\[
\frac{\Delta^i_{H(i-1) = h}}{E[\Delta^{i+1}_{H(i)}|H(i-1) = h]} = \frac{\alpha' + \beta' + 1}{\alpha' + \beta'}
\]

Since the prior \(p\) follows \(Be(\alpha, \beta)\), for any history \(h\) of first \(i - 1\) rounds, the distribution \(Be(\alpha', \beta')\) over \(p|H^{(i-1)} = h\) satisfies that \(\alpha' + \beta' = i + \alpha + \beta - 1\). So starting from any history of first \(i - 1\) rounds,

\[
\frac{\Delta^i_{H(i-1) = h}}{E[\Delta^{i+1}_{H(i)}|H(i-1) = h]} = \frac{i + \alpha + \beta}{i + \alpha + \beta - 1}
\]  

(5)

which only depends on \(i\), given fixed \(\alpha, \beta\).

Therefore

\[
\frac{E[\Delta^i_G]}{E[\Delta^{i+1}_G]} = \frac{E[\Delta^i_{H(i-1)}]}{E[\Delta^{i+1}_{H(i-1)}]}
\]

\[
= \frac{E[\Delta^i_{H(i-1)}]}{E_{H(i)}[E_{H(i)}[\Delta^{i+1}_{H(i)}|H(i-1)]]}
\]

\[
= \frac{E_{H(i-1)}[\frac{i + \alpha + \beta}{i + \alpha + \beta - 1} \cdot E_{H(i)}[\Delta^{i+1}_{H(i)}|H(i-1)]]}{E_{H(i-1)}[E_{H(i)}[\Delta^{i+1}_{H(i)}|H(i-1)]]}
\]

(\text{due to formula (5)})

\[
= \frac{i + \alpha + \beta}{i + \alpha + \beta - 1}
\]
Then we can use the expected surprise in round \( n - 1 \) to represent the amount of surprise in any round \( i \leq n - 1 \):

\[
\frac{E[\Delta^i_B]}{E[\Delta^{n-1}_B]} = \prod_{j=1}^{i-2} \frac{E[\Delta^j_B]}{E[\Delta^{j+1}_B]}
\]

\[
E[\Delta^i_B] = \frac{n + \alpha + \beta - 2}{i + \alpha + \beta - 1} \times E[\Delta^{n-1}_B]
\]

Therefore, the sum of the expected surprise in the first \( n - 1 \) rounds is

\[
\sum_{i=1}^{n-1} E[\Delta^i_B] = E[\Delta^{n-1}_B] \times \left( \sum_{i=1}^{n-1} \frac{n + \alpha + \beta - 2}{i + \alpha + \beta - 1} \right)
\]

\[
= E[\Delta^{n-1}_B] \times (n + \alpha + \beta - 2) \times H_{\alpha+\beta}(n - 1)
\]

\[\square\]

4 Finite Case

In this section, we follow our method overview to study the finite case. We first present our results in Section 4.1. We then show a general analysis in Section 4.2 and apply the results of the general analysis to study two special cases: 1) symmetric case: \( \alpha = \beta \) including the uniform case: \( \alpha = \beta = 1 \); 2) certain case \( \alpha = \lambda p, \beta = \lambda (1 - p), \lambda \to \infty \). Finally, we apply the results in Section 4.2 to show a linear algorithm for general beta prior setting.

4.1 Results in Finite Case

Recall that \( \text{RD}(x) := \text{the nearest integer to } x \text{ that has the same parity as } n \), and \( H := H_{\alpha+\beta}(n - 1) = \sum_{i=1}^{n-1} \frac{1}{i + \alpha + \beta - 1} \).

**Theorem 4.1.** For all \( \alpha \geq \beta \geq 1 \) \( n > 1 \),

- **Symmetric** \( \alpha = \beta \)

\[
x^*(\alpha, \alpha, n) = \text{RD}\left( \frac{n - 1}{2\alpha H - \frac{n - 1}{n + 2\alpha - 1}} \right)
\]

- **Uniform** \( \alpha = \beta = 1 \)

\[
x^*(1, 1, n) = \text{RD}\left( \frac{n - 1}{2H - \frac{n - 1}{n + 1}} \right)
\]

\[\text{When there is a tie, we pick the smaller one.}\]

\[\text{Note that assuming } \alpha \geq \beta \text{ does not lose generality since we can exchange Alice and Bob.}\]
• **Certain** \( \alpha = \lambda p, \beta = \lambda(1 - p), \lambda \to \infty \) Let \( F(x) := (2np - n - (x - 1))p^{x-1} + (n - 2np - (x - 1))(1-p)^{x-1}, \ x \in [1, n-1], \) \( F(x) = 0 \) has a trivial solution at \( x = 1 \) and a unique non-trivial solution \( \tilde{x} \) when \( p > \frac{1}{2} \) and \( n > \frac{1}{(\frac{1}{2} - p)\ln(\frac{1-p}{p})} \),

\[
x^*(\alpha, \beta, n) = \begin{cases} \text{RD}(\tilde{x}) & \text{if } p > \frac{1}{2} \text{ and } n > \frac{1}{(\frac{1}{2} - p)\ln(\frac{1-p}{p})} \\ \text{RD}(1) & \text{otherwise} \end{cases}
\]

Moreover, if \( p > \frac{\frac{1}{2} - p}{1 + (a+1) - \frac{1}{a}} \) where \( a = 2np - n - 2 > 1 \), then

\[
x^*(\alpha, \beta, n) \in [\text{RD}(2np - n) - 2, \text{RD}(2np - n) + 2],
\]

that is, the optimal bonus is around the “expected lead”.

• **General** There exists an \( O(n) \) algorithm to compute the optimal bonus \( x^*(\alpha, \beta, n) \).

We then present corresponding numerical results here. Based on Theorem 4.1, we draw the contours of \( \tilde{x} \) for varies of cases. The optimal \( x^* \) is \( \text{RD}(\tilde{x}) \). Though \( n \) can only be positive integers, we also smooth the contours for other \( n \). In the symmetric case, the optimal bonus size increases as the number of rounds \( n \) increases and the amount of uncertainty \( \frac{1}{2\alpha} \) decreases. In the certain case, as we predicted in theory, the optimal bonus size is close to “expected lead”. In the general case, as \( n \) gets larger, the result becomes closer to the asymptotic case (see Figure 1).

Finally, we provide additional numerical results illustrating how the overall surprise depends on the bonus size in Figure 9. For different settings, for all \( x \), we directly compute \( E[\Delta B(x)] \) by using backward induction to compute all belief curves. We also annotate our theoretical optimal bonus \( x^* = \text{RD}(\tilde{x}) \) based on Theorem 4.1. The overall surprise varies with the bonus size and in some cases (e.g. certain, \( n = 20, p = 0.7 \)), the optimal bonus creates surprise that doubles the amount of surprise created by the trivial settings (\( x = \text{RD}(0) \) or \( x = \text{RD}(n) \)). Moreover, the optimal bonus depends on the properties of the setting: number of rounds and uniform, symmetric, or skewed. Additionally, in the figures, we see that the curves all have a single peak, so the local and global optima coincide.

![Figure 6: Symmetric case: Optimal bonus \( x^* = \text{RD}(\tilde{x}) \)](image_url)

\[
\frac{1}{1 + (a+1) - \frac{1}{a}} < \frac{1}{1 + e^{-1}} \text{ and when } a \to +\infty, \frac{1}{1 + (a+1) - \frac{1}{a}} \to \frac{1}{2}
\]
(a) Optimal bonus $x^* = \text{RD}(\bar{x})$

(b) $2np - n$

Figure 7: Certain case

(a) $n = 5$

(b) $n = 10$

(c) $n = 20$

(d) $n = 40$

Figure 8: General case: Optimal bonus $x^* = \text{RD}(\bar{x})$. Here each area between two contour lines has the same optimal bonus $x^*$. For example, in $n = 5$, the red area’s optimal bonus size is 5, yellow is 3, cyan is 1.
Figure 9: Relation between bonus size and overall surprise

4.2 General Analysis

In this subsection, we derive a formula which can be applied to all settings in the later sections.

\[
E[\Delta_B(x)] = E[\Delta_B^{n-1}(x)] \ast (n + \alpha + \beta - 2) \ast \mathbb{H} + E[\Delta_B^n(x)]
\]  

(Lemma 3.1)

\[
= \left( \Pr[S_{n-2} = L_{n-2}] \ast \Delta_B^{n-1}_{S_{n-2} = L_{n-2}} + \Pr[S_{n-2} = U_{n-2}] \ast \Delta_B^{n-1}_{S_{n-2} = U_{n-2}} \right) 
\]

penultimate round (at point \( L_{n-2} \))  

\[
+ \sum_{j=L_{n-2}+1}^{U_{n-2}} \Pr[S_{n-2} = j] \ast \Delta_B^{n-1}_{S_{n-2} = j}
\]

penultimate round (between \( L_{n-2} \) and \( U_{n-2} \))

\[
+ \sum_{j=L_{n-1}}^{U_{n-1}} \Pr[S_{n-1} = j] \ast \Delta_B^n_{S_{n-1} = j}
\]

final round  

(recall (1) and (2) in method overview)
Here are shorthand notations:

\[
\begin{align*}
q^i_j &= \mathbb{E}[p|(S_i = j)], & 0 \leq i \leq n-1 \\
b^i_j &= \Pr(O = 1|(S_i = j)), & 0 \leq i \leq n-1 \\
d^i_{j+1} &= b^i_{j+1} - b^{i+1}_{j+1}, & 0 \leq i \leq n-2
\end{align*}
\]

The above definition for \(d^i_{j+1}\) is only for \(0 \leq i \leq n-2\) since \(d^{n-1}_n\) involves \(b^n_j\) but the definition for \(b^i_j\) is for \(0 \leq i \leq n-1\). The final round’s belief value is either 0 or 1 and depends on both the number of rounds Alice wins among the first \(n-1\) rounds \((S_{n-1})\) and whether Alice wins the final round \((H_n = +\) or \(H_n = -\)). Thus, we define the belief change in the final round directly as follows.

\[
d^{n-1}_j := \Pr(O = 1|(H_n = +) \land (S_{n-1} = j)) - \Pr(O = 1|(H_n = -) \land (S_{n-1} = j))
\]

In fact, for the no-surprise red/blue points in the final round (Figure 3), the belief change is 0, for other grey points, the belief change is 1.

By substituting the above shorthand, we have
Lemma 4.2. We have

\[
q_j^i = \frac{j + \alpha}{i + \alpha + \beta} \\
Pr[S_i = j] = \frac{(\alpha + \beta - 1)\binom{j}{\alpha + \beta - 2}}{\binom{\alpha + \beta + i - 1}{j + \alpha - 1}}
\]

\[
Q_j^i = \frac{2(\alpha + \beta - 1)\binom{i}{\alpha + \beta - 2}}{\binom{\alpha + i + \beta}{j + \alpha}}
\]

and

\[
d_j^{n-2} = \begin{cases} 
0, & j < L_{n-2} \\
\frac{n-1 + \alpha}{n + \alpha + \beta - 1}, & j = L_{n-2} \\
\frac{n + 1}{n + \alpha + \beta - 1} - \frac{n + \alpha + \beta - 1}{n + \alpha + \beta - 1}, & L_{n-2} < j < U_{n-2} \\
1, & j = U_{n-2} \\
0, & j > U_{n-2}
\end{cases}
\]

\[
d_j^{n-1} = \begin{cases} 
0, & j < L_{n-1} \\
1, & L_{n-1} \leq j \leq U_{n-1} \\
0, & j > U_{n-1}
\end{cases}
\]

The expected surprisal in a final round is

\[
E[\Delta_B(x)] = \left( \Pr[S_{n-2} = L_{n-2}] * 2q_{L_{n-2}}^{n-2} * (1 - q_{L_{n-2}}^{n-2}) * d_{L_{n-2}}^{n-2} \right)
\]

penultimate round (at point \(L_{n-2}\))

\[
+ \Pr[S_{n-2} = U_{n-2}] * 2q_{U_{n-2}}^{n-2} * (1 - q_{U_{n-2}}^{n-2}) * d_{U_{n-2}}^{n-2}
\]

penultimate round (at point \(U_{n-2}\))

\[
+ \sum_{j=L_{n-2}+1}^{U_{n-2}-1} \Pr[S_{n-2} = j] * 2q_j^{n-2} * (1 - q_j^{n-2}) * d_j^{n-2} \right) * (n + \alpha + \beta - 2) * \mathbb{H} 
\]

penultimate round (between \(L_{n-2}\) and \(U_{n-2}\))

\[
+ \sum_{j=L_{n-1}}^{U_{n-1}-1} \Pr[S_{n-1} = j] * 2q_j^{n-1} * (1 - q_j^{n-1}) * d_j^{n-1}
\]

final round

\[
(6)
\]

We further introduce a shorthand \(Q_j^i := \Pr[S_i = j] * 2q_j^i * (1 - q_j^i), 0 \leq i \leq n - 1\) and by substituting this shorthand, we have

\[
E[\Delta_B(x)] = \left( Q_{L_{n-2}}^{n-2} * d_{L_{n-2}}^{n-2} + Q_{U_{n-2}}^{n-2} * d_{U_{n-2}}^{n-2} \right)
\]

penultimate round (at point \(L_{n-2}\))

\[
+ \sum_{j=L_{n-2}+1}^{U_{n-2}-1} Q_j^{n-2} * d_j^{n-2} \right) * (n + \alpha + \beta - 2) * \mathbb{H} + \sum_{j=L_{n-1}}^{U_{n-1}-1} Q_j^{n-1} * d_j^{n-1}
\]

penultimate round (between \(L_{n-2}\) and \(U_{n-2}\))

As we mentioned in the overview, we pick the final round and the penultimate round to represent the overall expected surprisal since the belief change, in these two rounds, has a simple representation, as stated in the following lemma.

Lemma 4.2. We have

\[16\text{When } \alpha + \beta \text{ is not an integer, we use } \binom{n}{k} := \frac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k+1)} \text{ as the continuous generalization.} \]
We prove the lemma by delicate analysis based on the properties of Beta distribution. We defer the proof to appendix. Based on Lemma 4.2, we substitute the final two rounds’ belief change \( d_j^{n-1}, d_j^{n-2} \) into formula (7):

\[
E[\Delta_B(x)] = \begin{cases} 
\text{penultimate round (at point } L_{n-2} \text{)} & + \\
\begin{aligned}
Q_{n-2}^{L_{n-2}} & \frac{L_{n-1} + \alpha}{n + \alpha + \beta - 1} \\
+ & \sum_{j=L_{n-2}+1}^{n-2-1} Q_{n-2}^j \frac{1}{n + \alpha + \beta - 1}
\end{aligned} & \text{penultimate round (at point } U_{n-2} \text{)} \\
\text{final round} & + \\
\begin{aligned}
Q_{n-2}^{n-2} & \frac{n-x+1/2}{n + \alpha + \beta - 1} + Q_{n-2}^{U_{n-2}} \frac{n-x+1}{n + \alpha + \beta - 1} \\
+ & \sum_{j=\frac{n-x}{2}}^{\frac{n+x}{2}} Q_{n-1}^j \frac{1}{n + \alpha + \beta - 1}
\end{aligned}
\end{cases} 
\]  \( E[\Delta_B(x)] \) (8)  

In order to find the optimal \( x \), we calculate

\[
E[\Delta_B(x+1)] - E[\Delta_B(x-1)] = \begin{cases} 
\text{penultimate round (at point } L_{n-2} \text{)} & + \\
\begin{aligned}
Q_{n-2}^{L_{n-2}} & \frac{L_{n-1} + \alpha}{n + \alpha + \beta - 1} \\
+ & \sum_{j=L_{n-2}+1}^{U_{n-1}} Q_{n-2}^j \frac{1}{n + \alpha + \beta - 1}
\end{aligned} & \text{penultimate round (at point } U_{n-2} \text{)} \\
\text{final round} & + \\
\begin{aligned}
Q_{n-2}^{n-2} & \frac{n-x+1}{n + \alpha + \beta - 1} + Q_{n-2}^{U_{n-2}} \frac{n-x+1/2}{n + \alpha + \beta - 1} \\
+ & \sum_{j=\frac{n-x}{2}}^{\frac{n+x}{2}} Q_{n-1}^j \frac{1}{n + \alpha + \beta - 1}
\end{aligned}
\end{cases} 
\]

Then the following claim shows that we can obtain the optimal bonus by finding “local maximum” \( \tilde{x} \). We defer the proof to appendix.

**Claim 4.3** (Local Maximum \( \rightarrow \) Optimal Bonus). *If there exists \( \tilde{x} \in (0, n+1) \) such that for all \( 1 \leq x < \tilde{x} \), \( E[\Delta_B(x+1)] \geq E[\Delta_B(x-1)] \) and when \( \tilde{x} \leq n-1 \), for all \( \tilde{x} \leq x \leq n-1 \), \( E[\Delta_B(x+1)] \leq E[\Delta_B(x-1)] \), then \( RD(\tilde{x}) \) is the optimal bonus.*

Later we show that formula (9) induces a linear algorithm to find the optimal bonus size in general and can be significantly simplified in the symmetric case and certain case.

\footnote{For \( x < 2 \), we have \( L_{n-2} + 1 > U_{n-2} - 1 \), and for \( x = 0 \), we have \( L_{n-1} = U_{n-1} \). We define the summation from a larger subscript to a smaller superscript as zero. This definition is valid since in those cases, no surprise is generated.}
4.3 Symmetric Case

We start to analyze the symmetric case.

Observation 4.4. In the symmetric case, i.e. $\alpha = \beta$, $Q^i_j$ is also symmetric, that is $Q^i_j = Q^i_{i-j}$.

We defer the proof to appendix. Based on the above observation, we can further simplify the formula (9)

$$E[\Delta_B(x + 1)] - E[\Delta_B(x - 1)]$$

$$= \left( \frac{n-2}{n-x-3} - \frac{n-2}{n-x-1} \right) \ast \left( \frac{n-x-1}{2} + \alpha \right) + \left( \frac{n-2}{n-x+1} - \frac{n-2}{n-x-3} \right) \ast \left( \frac{n-x-1}{2} + \alpha \right)$$

(based on formula (9))

$$= \left( \frac{n-2}{n-x-3} - \frac{n-2}{n-x-1} \right) \ast \left( \frac{n-x-1}{2} + \alpha \right) + \left( \frac{n-2}{n-x+1} - \frac{n-2}{n-x-3} \right) \ast \left( \frac{n-x-1}{2} + \alpha \right)$$

$$\left( Q^i_{n-x-1} = Q^i_{n-x+1} = Q^i_{n-x-3} = Q^i_{n-x-1} \right)$$

(according to Observation 4.4)

$$= \left( \frac{n-2}{n-x-3} - \frac{n-2}{n-x-1} \right) \ast \left( n-x-1 + 2\alpha \right) \ast \frac{(n+2\alpha-2)}{(n+2\alpha-1)} + 2Q^i_{n-2-x-1}$$

(10)

The remaining task is to analyze $Q^i_j$'s in the above formula. They share some components which can be used for further simplification. We start from the uniform case and analyze it step by step.

Uniform Case. In this case $\alpha = \beta = 1$ and we can substitute $\alpha, \beta$ in Lemma 4.2 and obtains that the probability that Alice wins any number of rounds in the first $i$ rounds is equal, that is

$$\Pr[S_i = j] = \frac{1}{i+1}$$

and $Q^i_j$ is

$$Q^i_j = \frac{2(j+1)(i+1-j)}{(i+1)(i+2)^2}$$

We substitute $Q^i_j$ into formula (10)
\[ E[\Delta_B(x + 1)] - E[\Delta_B(x - 1)] \]
\[ = \left( Q^{n-2}_{\frac{n-1}{2}} - Q^{n-2}_{\frac{n-4}{2}} \right) \ast (n - x - 1) + 2 \ast \frac{(n + 2\alpha - 2)\mathbb{H}}{n + 2\alpha - 1} + 2Q^{n-1}_{\frac{n-1}{2}} \]  
(based on formula (10))
\[ = \left( \frac{(n - x - 1)(n + x + 1)}{2(n - 1)n^2} - \frac{(n - x + 1)(n + x - 1)}{2(n - 1)n^2} \right) \ast (n - x + 1) \ast \frac{n\mathbb{H}}{n + 1} \]
\[ + \frac{(n - x + 1)(n + x - 1)}{n(n + 1)^2} \]
\[ = \frac{(-4x) \ast (n - x + 1)}{2(n - 1)n^2} \ast \frac{n\mathbb{H}}{n + 1} + \frac{(n - x + 1)(n + x - 1)}{n(n + 1)^2} \]
\[ = \frac{(n - x + 1)((n - 1 - 2(1 + n)\mathbb{H})x + n^2 - 1)}{(n - 1)n(n + 1)^2} \]

To find \( \tilde{x} \) that satisfies conditions in Claim 4.3, we solve the equation
\[ \frac{(n - x + 1)((n - 1 - 2(1 + n)\mathbb{H})x + n^2 - 1)}{(n - 1)n(n + 1)^2} = 0, \]
and get
\[ x = \begin{cases} 
\frac{n^2 - 1}{2(1 + n)\mathbb{H} - n + 1} \\
n + 1 
\end{cases}. \]

Recall that \( x \leq n \), we discard the solution of \( x = n + 1 \), and then pick \( \tilde{x} := \frac{n^2 - 1}{2(1 + n)\mathbb{H} - n + 1} = \frac{n - 1}{2\mathbb{H} - \frac{1}{n + 1}} \).

When \( x < \tilde{x} \), we have \( (n - 1 - 2(1 + n)\mathbb{H})x + n^2 - 1 > 0 \), thus \( E[\Delta_B(x + 1)] - E[\Delta_B(x - 1)] > 0 \). Otherwise, \( E[\Delta_B(x + 1)] - E[\Delta_B(x - 1)] \leq 0 \). Based on Claim 4.3, the optimal bonus is
\[ x^*(1, 1, n) = \text{RD}\left( \frac{n - 1}{2\mathbb{H} - \frac{1}{n + 1}} \right). \]

**Symmetric Case**  Lemma 4.2 shows that if the prior \( p \) follows \( B\epsilon(\alpha, \alpha) \), then the probability of Alice wins \( j \) rounds in the first \( i \) rounds is
\[ \Pr[S_i = j] = \frac{(2\alpha - 1)(\binom{j}{i}) (2\alpha - 2)}{(2\alpha + i - 1)(\binom{i + 2\alpha - 2}{j + \alpha - 1})} \]
and \( Q_j^i \) is
\[ Q_j^i = \frac{2(2\alpha - 1)(\binom{j}{i}) (2\alpha - 2)}{(2\alpha + i)(\binom{i + 2\alpha}{j + \alpha})} \]

Then we substitute \( Q_j^i \) into formula (10).
4.4 Certain Case

In the certain case, \( \alpha = \lambda p, \beta = \lambda(1 - p), \lambda \to \infty \). Thus, when \( n \) is finite, the winning probability of Alice is fixed to \( p \) across all rounds. Note that we only consider \( \alpha \geq \beta \) without...
loss of generality. Therefore, \( p \geq \frac{1}{2} \). The number of rounds Alice wins follows a binomial distribution. Formally, the probability of Alice wins \( j \) rounds in the first \( i \) rounds is

\[
\Pr[S_i = j] = \binom{i}{j} p^j (1 - p)^{i-j}
\]

Then we calculate \( Q_j^i \)

\[
Q_j^i = 2^i \binom{i}{j} p^{i+1} (1 - p)^{i-j+1} \tag{11}
\]

Moreover, we can apply the Main Technical Lemma 3.1 to show that the first \( n - 1 \) rounds have the same expected surprise.

**Corollary 4.5.** When \( \alpha = \lambda p, \beta = \lambda (1 - p) \), \( \lambda \to \infty \), given \( n \) that is finite,

\[
E[\Delta_B] = \sum_{i=1}^{n} E[\Delta_B^i] = E[\Delta_B^{n-1}] \ast (n - 1) + E[\Delta_B^n]
\]

**Proof of Corollary 4.5**

\[
\frac{E[\Delta_B^i]}{E[\Delta_B^{i+1}]} = \frac{i + \alpha + \beta}{i + \alpha + \beta - 1} \to 1
\]

as \( \lambda \to \infty \).

By formula (8), we have

\[
E[\Delta_B(x)] = \left( Q_{n-x-2}^{n-2} \ast p + Q_{n-x-2}^{n-2} \ast (1 - p) \right) \ast (n - 1) + \sum_{j = \frac{n-x}{2}}^{\frac{n+x-2}{2}} Q_j^{n-1}
\]

(Formula (8) and Corollary 4.5)

\[
= \left( Q_{n-x-2}^{n-2} \ast p + Q_{n-x-2}^{n-2} \ast (1 - p) \right) \ast (n - 1) + \sum_{j = \frac{n-x}{2}}^{\frac{n+x-2}{2}} Q_j^{n-1}
\]

\[
= \left( 2^\left(\frac{n-2}{2}\right)p^{\frac{n-x+2}{2}} (1 - p)^{\frac{n+x}{2}} + 2^\left(\frac{n-2}{2}\right)p^{\frac{n+x}{2}} (1 - p)^{\frac{n-x+4}{2}} \right) \ast (n - 1)
\]

\[
+ \sum_{j = \frac{n-x}{2}}^{\frac{n+x-2}{2}} 2^\left(\frac{n-1}{j}\right)p^{j+1} (1 - p)^{n-j} \quad \text{(Apply formula (11) to substitute } Q_j^i)\]

Then we calculate
\[
E[\Delta_B(x + 1)] - E[\Delta_B(x - 1)]
\]
\[
= 2(n - 1) \left( \frac{n - 2}{n - x - 3} \right) \left( \frac{n - x + 1}{2} \right) p \frac{n - x + 1}{2} (1 - p) + \left( \frac{n - 2}{n + x - 1} \right) p \frac{n + x + 1}{2} (1 - p)
\]

\[
\text{difference of first n-1 round’s expected surprise}
\]
\[
- \left( \frac{n - 2}{n - x - 3} \right) p \frac{n - x + 1}{2} (1 - p) \frac{n - x + 1}{2} - \left( \frac{n - 2}{n + x - 1} \right) p \frac{n + x + 1}{2} (1 - p) \frac{n + x + 3}{2}
\]

\[
\text{difference of final round’s expected surprise}
\]
\[
+ \sum_{i = \frac{n - 1}{2}}^{\frac{n + x - 1}{2}} 2^{(n - 1)} \left( \frac{n - i}{i} \right)^{p+1}(1 - p)^{n-i} - \sum_{i = \frac{n - 1}{2}}^{\frac{n + x - 3}{2}} 2^{(n - 1)} \left( \frac{n - i}{i} \right)^{p+1}(1 - p)^{n-i}
\]

\[
= (n - x - 1) \left( \frac{n - 1}{n - x - 3} \right) p \frac{n - x + 1}{2} (1 - p) \frac{n - x + 1}{2} + \left( \frac{n - 1}{n + x - 1} \right) p \frac{n + x + 1}{2} (1 - p) \frac{n + x + 3}{2}
\]
\[
- (n + x - 1) \left( \frac{n - 1}{n - x - 3} \right) p \frac{n - x + 1}{2} (1 - p) \frac{n - x + 1}{2} + \left( \frac{n - 1}{n + x - 1} \right) p \frac{n + x + 1}{2} (1 - p) \frac{n + x + 3}{2}
\]
\[
+ 2 \left( \frac{n - 1}{n - x - 1} \right) p \frac{n + x + 1}{2} (1 - p) \frac{n + x + 1}{2} + 2 \left( \frac{n - 1}{n + x - 1} \right) p \frac{n + x + 1}{2} (1 - p) \frac{n + x + 1}{2}
\]

(Apply \( n(n-1) = (n-k)k \))

\[
= \left( \frac{n - 1}{n + x - 1} \right) p \frac{n + x + 1}{2} (1 - p) \frac{n + x + 1}{2} \left( (n - x + 1)((1-p)x + p^x) - (n + x - 1)(p(1-p)^{x-1} + p^{x-1}(1-p)) \right)
\]

\[
= \left( \frac{n - 1}{n + x - 1} \right) p \frac{n + x + 1}{2} (1 - p) \frac{n + x + 1}{2} \left( (2np - n - (x - 1))p^{x-1} + (n - 2np - (x - 1))(1-p)^{x-1} \right)
\]

\( \propto (2np - n - (x - 1))p^{x-1} + (n - 2np - (x - 1))(1-p)^{x-1} \)

(Since \( \left( \frac{n - 1}{n + x - 1} \right) p \frac{n + x + 1}{2} (1 - p) \frac{n + x + 1}{2} > 0 \))

We define \( F(x) := (2np - n - (x - 1))p^{x-1} + (n - 2np - (x - 1))(1-p)^{x-1}, x \in [1, n + 1] \) and analyze the solution of \( F(x) = 0 \). We show several examples of \( F(x) \) with different \( n \) in Figure [11]. We can see that, intuitively, when \( n \) is large enough, \( F(x) = 0 \) has 2 solutions. One is \( x = 1 \), and the other solution is close to \( 2np - n \) when \( n \) is large.

![Figure 11: Examples for function F](image-url)

(a) \( n = 5 \)  
(b) \( n = 6 \)  
(c) \( n = 7 \)  
(d) \( n = 8 \)
Lemma 4.6. When $p \geq \frac{1}{2}$, $F(x) = 0$ has a trivial solution at $x = 1$, and has a non-trivial solution $\tilde{x} \in (1, 2np - n + 1)$ if and only if $p > \frac{1}{2}$ and $n > \frac{1}{(\frac{1}{2} - p)\ln(\frac{1}{p})}$. There is no other solution. Moreover, when $x \in (1, \tilde{x})$, $F(x) > 0$, when $\tilde{x} < n - 1$ and $x \in (\tilde{x}, n - 1)$, $F(x) < 0$. Besides, let $a = 2np - n - 2$, when $p > \frac{1}{1+(a+1)^{-\frac{1}{2}}}$, the non-trivial solution of $F(x) = 0$ is in $(2np - n - 1, 2np - n + 1)$.

We defer the proof to appendix. $F(x)$ has a solution $\tilde{x}$ in $(1, 2np - n + 1)$ when $p > \frac{1}{2}$ and $n > \frac{1}{(\frac{1}{2} - p)\ln(\frac{1}{p})}$. When $x \in [1, \tilde{x})$, $F(x) \geq 0$, i.e. $E[\Delta_B(x + 1)] - E[\Delta_B(x - 1)] \geq 0$. When $\tilde{x} \leq n - 1$ and $x \in [\tilde{x}, n - 1)$, $E[\Delta_B(x + 1)] - E[\Delta_B(x - 1)] \leq 0$. Therefore, the conditions of Claim 4.3 are satisfied. We apply Claim 4.3 to obtain:

$$x^*(\alpha, \beta, n) = \begin{cases} 
\text{RD}(\tilde{x}) & \text{if } p > \frac{1}{2} \text{ and } n > \frac{1}{(\frac{1}{2} - p)\ln(\frac{1}{p})} \\
\text{RD}(1) & \text{otherwise}
\end{cases}$$

where $\tilde{x}$ is the non-trivial solution of the equation $F(x) = 0$.

Finally, we study the approximation of the $x^*(\alpha, \beta, n)$ in the certain case and show that under certain conditions, it is around the "expected lead" $\text{RD}(2np - n)$, the number of points the weaker player needs to comeback in expectation.

Lemma 4.6 shows that when $p > \frac{1}{1+(a+1)^{-\frac{1}{2}}}$, the non-trivial solution of $F(x) = 0$ is in $(2np - n - 1, 2np - n + 1)$. Recall that $\text{RD}(x) \in [x - 1, x + 1)$. Then we have $|\text{RD}(\tilde{x}) - \text{RD}(2np - n)| < 3$. Moreover, $\text{RD}(x)$ is an integer that has the same parity as $n$. Therefore, $|\text{RD}(\tilde{x}) - \text{RD}(2np - n)| \leq 2$, i.e. the difference between the approximation $\text{RD}(2np - n)$ and the optimal bonus $\text{RD}(\tilde{x})$ is $\leq 2$.

4.5 General Beta Prior Setting

We provide an $O(n)$ algorithm for general Beta prior setting. A natural idea is to enumerate all possible bonus $x$ and calculate the corresponding expected total surprise value $E[\Delta_B(x)]$. However, $E[\Delta_B(x)]$ contains an $O(n)$ summation and lead to an $O(n^3)$ running time algorithm.

Recall formula [9].

$$E[\Delta_B(x + 1)] - E[\Delta_B(x - 1)] = \left(Q_{\frac{n-2}{2}+1}^{n-2} - Q_{\frac{n-2}{2}}^{n-2}\right) \ast \left(\frac{n-x-1}{2} + \alpha\right) + \left(Q_{\frac{n-2}{2}+1}^{n-2} - Q_{\frac{n-2}{2}}^{n-2}\right) \ast \left(\frac{n-x-1}{2} + \beta\right) * \frac{(n + 2\alpha - 2)\mathbb{H}}{n + 2\alpha - 1} + Q_{\frac{n-1}{2}+1}^{n-1} + Q_{\frac{n-1}{2}+1}^{n-1}$$

Note that in this formula, only $\mathbb{H}$ and $Q^i_j$ cannot be calculated in $O(1)$ time. We can preprocess $\mathbb{H}$ in the time of $O(n)$. Then notice that if we can calculate all the $Q^i_j$ in the formula within $O(1)$ time, we can compute the difference between $E[\Delta_B(x + 1)]$ and $E[\Delta_B(x - 1)]$ in $O(1)$ time.

Recall Lemma 4.2.
The complexity of the algorithm is optimal.

We implement the algorithm to conduct numerical experiments for $n = 5, n = 10, n = 20, n = 40$ (Figure 5) and the figure becomes more and more similar to the asymptotic case (Figure 1) as $n$ increases.
5 Asymptotic Case

In the asymptotic case, we can use a continuous integral to approximate the discrete summation. Here we define bonus ratio \( \mu := \frac{2}{n} \) and use the integration of \( \mu \) to approximate the overall surprise. Formally,

Theorem 5.1. For all \( \alpha \geq \beta \geq 1 \), there exists a function \( Z_{\alpha, \beta, n}(\mu) \) such that \( \forall \mu \in (0, 1), \ E[\Delta S(\mu * n)] = Z_{\alpha, \beta, n}(\mu) * (1 + O(\frac{1}{n})) \). When we define \( \mu^* := \arg \max_{\mu} Z_{\alpha, \beta, n}(\mu) \),

- **Symmetric** \( \alpha = \beta \)
  \[
  \mu^* = \frac{1}{2\alpha_{\beta,n}(n-1) - 1}
  \]

- **Near-certain** \( \alpha = \lambda p, \beta = \lambda (1 - p) \) fixing \( p \), for all sufficiently small \( \epsilon > 0 \), when \( \lambda > O(\log \frac{1}{\epsilon}) \) the optimal \( \mu^* \) is around the “expected lead”,
  \[
  \mu^* \in \left( \frac{\alpha - \beta}{\alpha + \beta} \right) H + 1 - \epsilon, \left( \frac{\alpha - \beta}{\alpha + \beta} \right) H - 1 = \left( \frac{\alpha - \beta}{\alpha + \beta} - \epsilon, \frac{\alpha - \beta}{\alpha + \beta} \right) = (2p - 1 - \epsilon, 2p - 1)
  \]

- **General** \( \mu^* \) is the unique solution of \( G(\mu) = 0 \) and \( \mu^* < \frac{(\alpha - \beta) H + 1}{(\alpha + \beta) H - 1} \).
  \[
  G(\mu) := (1 + \mu)^{\alpha - \beta} \left( \frac{\alpha - \beta}{\alpha + \beta} \right) H + 1 - \mu + (1 - \mu)^{\alpha - \beta} \left( \frac{\alpha - \beta}{\alpha + \beta} \right) H - 1 - \mu
  \]

![Figure 12: Examples for function G](image)

We define \( Beta(\theta; \alpha, \beta) \) as the density function of Beta distribution, i.e. \( Beta(\theta; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1} \). In the asymptotic case when \( n \) is sufficiently large, we can simplify Lemma 4.2 to obtain Lemma 5.2 illustrated in Figure 13. The simplification for \( d, L, U \) is straightforward. For \( Pr[S_{n-1} = j] \), note that informally, due to law of large number, when \( n \) is sufficiently large, the state in the \( n - 1 \) or \( n - 2 \) round concentrate to \( pn \). Thus, given that \( p \) follows distribution \( Beta(\theta; \alpha, \beta) \), we have \( Pr[S_{n-1} = j] \approx Pr[S_{n-2} = j] \) which are approximately proportional to \( Beta(\theta_j; \alpha, \beta) \), where \( \theta_j = \frac{\alpha}{\beta} \).
Lemma 5.2 (Informal). When $n$ is sufficiently large, fixed $\alpha, \beta, \mu \in (0,1)$, we have

$$\begin{cases}
\frac{L_{n-2}}{n} & \approx \frac{L_{n-1}}{n} & \approx \frac{1-\mu}{2} \\
\frac{U_{n-2}}{n} & \approx \frac{U_{n-1}}{n} & \approx \frac{1+\mu}{2}
\end{cases}$$

For any $L_{n-1} \leq j \leq U_{n-1}$, let $\theta_j = \frac{j}{n}$, we have

$$\Pr[S_{n-1} = j] \approx \Pr[S_{n-2} = j] \approx \frac{\text{Beta}(\theta_j; \alpha, \beta)}{n}$$

and

$$d_j^{n-2} = \begin{cases}
0, & \theta_j < \frac{L_{n-2}}{n} \\
\frac{1-\mu}{2}, & \frac{L_{n-2}}{n} \leq \theta_j < \frac{U_{n-2}}{n} \\
\frac{1}{n}, & \frac{U_{n-2}}{n} \leq \theta_j < \frac{1-\mu}{2} \\
0, & \theta_j \geq \frac{1-\mu}{2}
\end{cases}$$

and

$$d_j^{n-1} = \begin{cases}
0, & \theta_j < \frac{L_{n-1}}{n} \\
\frac{1-\mu}{2}, & \frac{L_{n-1}}{n} \leq \theta_j < \frac{U_{n-1}}{n} \\
1, & \frac{U_{n-1}}{n} \leq \theta_j \\
0, & \theta_j > \frac{U_{n-1}}{n}
\end{cases}$$

Figure 13: Asymptotic case
By substituting Lemma 5.2 into formula (6), we can obtain an approximation for the overall expected surprise $E[\Delta_B(\mu \ast n)]$:

$$Z_{\alpha,\beta,n}(\mu) = \begin{align*}
&= \left(\frac{Beta\left(\frac{1}{2}; \alpha, \beta\right)}{n} \ast 2 \ast \frac{1 + \mu}{2} \ast \frac{1 - \mu}{2} + \frac{Beta\left(\frac{1 + n}{2}; \alpha, \beta\right)}{n} \ast 2 \ast \frac{1 + \mu}{2} \ast \frac{1 - \mu}{2} \right) \ast nH + \int_{\frac{1 + \mu}{2}}^{1} Beta(\theta; \alpha, \beta) \ast 2 \ast \theta \ast (1 - \theta) \ast \frac{1}{n} d\theta \\
&\quad + \left(\int_{\frac{1}{2}}^{\frac{1 + \mu}{2}} 2Beta(\theta; \alpha, \beta) \theta (1 - \theta) d\theta\right) \ast (H + 1) \\
&\quad + \left(\beta(\frac{1 + \mu}{2}; \alpha, \beta) + Beta\left(\frac{1}{2}; \alpha, \beta\right)\right) \left((1 - \mu)^2(1 + \mu)H + \frac{1}{4}\right)
\end{align*}
$$

Formally, we need a formal version of Lemma 5.2 to delicately analyze the relationship between $Z_{\alpha,\beta,n}(\mu)$ and $E[\Delta_B(\mu \ast n)]$ and prove the main result of Theorem 5.1. We defer the formal version of Lemma 5.2 and the proof to appendix.
In order to find the optimal $\mu$, we calculate the derivation of $Z_{\alpha,\beta,n}(\mu)$

$$\frac{dZ_{\alpha,\beta,n}(\mu)}{d\mu} = \text{Beta}(\frac{1+\mu}{2};\alpha,\beta) + \text{Beta}(\frac{1-\mu}{2};\alpha,\beta) \frac{1-\mu^2}{4} (H + 1)$$

$$+ \frac{d}{d\mu} \left( \text{Beta}(\frac{1-\mu}{2};\alpha,\beta) + \frac{(\alpha - \beta - (\alpha + \beta - 2)\mu) \text{Beta}(\frac{1+\mu}{2};\alpha,\beta) + (\alpha - \beta - (\alpha + \beta - 2)\mu) \text{Beta}(\frac{1-\mu}{2};\alpha,\beta)}{2} \right) (1 - \mu) H$$

$$= (1 - \mu) \text{Beta}(\frac{1+\mu}{2};\alpha,\beta) - (1 - \mu) \text{Beta}(\frac{1-\mu}{2};\alpha,\beta) \left( -((\alpha + \beta) H - 1) \mu + (\alpha - \beta) H + 1 \right)$$

$$(\text{Combining like terms})$$

$$+ (1 - \mu) \text{Beta}(\frac{1+\mu}{2};\alpha,\beta) - (1 - \mu) \text{Beta}(\frac{1-\mu}{2};\alpha,\beta) \left( -((\alpha + \beta) H - 1) \mu + (\alpha - \beta) H + 1 \right)$$

$$(\text{Substituting density of Beta distribution})$$

$$+ (1 - \mu) \text{Beta}(\frac{1+\mu}{2};\alpha,\beta) - (1 - \mu) \text{Beta}(\frac{1-\mu}{2};\alpha,\beta) \left( -((\alpha + \beta) H - 1) \mu + (\alpha - \beta) H + 1 \right)$$

$$((\mu < 1))$$

$$\alpha(1 + \mu)^{\alpha-1} - (1 - \mu)^{\beta-1} \left( -((\alpha + \beta) H - 1) \mu + (\alpha - \beta) H + 1 \right)$$

Let

$$G(\mu) := (1 + \mu)^{\alpha-\beta} \left( \frac{(\alpha - \beta) H + 1}{(\alpha + \beta) H - 1} - \mu \right) + (1 - \mu)^{\alpha-\beta} \left( \frac{-\alpha + \beta) H + 1}{(\alpha + \beta) H - 1} - \mu \right)$$

The examples of $G(\mu)$ are illustrated in Figure 12.

**Lemma 5.3** (Property of $G(\mu)$). For all $\alpha \geq \beta$, when $n$ is sufficiently large, $G(0) > 0$ and $G(1) < 0$. $G(\mu) = 0, \mu \in [0,1]$ has a unique solution and the solution is in $(0, (\alpha - \beta) H + 1)$. Moreover, for all $0 < \epsilon < \frac{(\alpha - \beta) H + 1}{(\alpha + \beta) H - 1}$, when $\alpha - \beta > \frac{\log(\frac{(\alpha - \beta) H + 1}{(\alpha + \beta) H - 1} - \epsilon) - \log \epsilon}{\log(1 + (\frac{(\alpha - \beta) H + 1}{(\alpha + \beta) H - 1} - \epsilon) - \log(1 - (\frac{(\alpha - \beta) H + 1}{(\alpha + \beta) H - 1} - \epsilon)))} \geq 20$, the solution is within $(\frac{(\alpha - \beta) H + 1 - \epsilon}{(\alpha + \beta) H - 1}, \frac{(\alpha - \beta) H + 1}{(\alpha + \beta) H - 1}) \approx (\frac{\alpha - \beta}{\alpha + \beta} - \epsilon, \frac{\alpha - \beta}{\alpha + \beta})$. 

\[20\text{This is approximately} \frac{\log \frac{\alpha - \beta}{\alpha + \beta} - \frac{1}{2}}{\log \alpha - \log \beta} < \frac{1}{\log \alpha - \log \beta} \frac{1}{\frac{1}{\epsilon}}\]
We defer the proof to appendix. The above lemma implies that $Z_{\alpha,\beta,n}(\mu)$ first increases and then decreases. Thus, the global optimal $\mu^*$ is also the local maximum, $G(\mu) = 0$'s unique solution. Thus, the results for general case in Theorem 5.1 follow from the lemma. In the symmetric case, $\alpha = \beta$,

$$G(\mu) = \frac{2}{(\alpha + \beta)H - 1} - 2\mu = 0$$

$$\Rightarrow \mu^* = \frac{1}{2\alpha H - 1}.$$

The results for near certain case also directly follow from the above lemma. Therefore, we finish the proof.

6 Conclusion and Discussion

In a multi-round competition, we show that we can increase the audience’s overall surprise by setting a proper bonus in the final round. We further show that the optimal bonus size depends on the audience’s prior and in the following settings, we obtain solutions of various forms for both the case of a finite number of rounds and the asymptotic case:

**Symmetric** the audience’s prior belief does not lean towards any player, here we obtain a clean closed-form solution in both the finite the and asymptotic case;

**Certain** the audience is a priori certain about the two players’ relative abilities, here the optimal bonus is a special function’s solution, and approximately and asymptotically equal to the “expected lead”, the number of points the weaker player will need to come back in expectation;

**General** the optimal bonus can be obtained by a linear algorithm and, in the asymptotic case, is a special function’s solution.

One natural extension is to validate our theoretical predictions using field experiments. We can potentially conduct online field experiments, e.g., AB test, to examine the effectiveness of the scoring rules. Moreover, the results from field experiments could potentially capture features that might be neglected in the existing models and consequently inform the development of new theories.

Regarding the theoretical work, one future direction is to incorporate the time factor into the model since a line of psychology literature shows that the audience may judge their experience largely on their feeling in the later part of the game (end-effect). We can generalize our analysis to factor in that surprise may be more valued in some time periods (e.g. the last round) than others. Another future direction is extending our results to the setting where we allow the scores to increase from the first round to the final round (e.g., 1,2,3,4,.....), or the bonus size depends on the results of previous rounds, or the setting where the game may not end in a pre-determined fixed time (e.g. Quidditch), or even a more complicated game setting (e.g. board games).
References

[1] Alan D Baddeley and Graham Hitch. The recency effect: Implicit learning with explicit retrieval? Memory & Cognition, 21(2):146–155, 1993.

[2] Paolo Bizzozero, Raphael Flepp, and Egon Franck. The importance of suspense and surprise in entertainment demand: Evidence from wimbledon. Journal of Economic Behavior & Organization, 130:47–63, 2016.

[3] Steven J Brams and Mehmet S Ismail. Making the rules of sports fairer. SIAM Review, 60(1):181–202, 2018.

[4] Mark Braverman, Omid Etesami, and Elchanan Mossel. Mafia: A theoretical study of players and coalitions in a partial information environment. The Annals of Applied Probability, 18(3):825–846, 2008.

[5] Babatunde Buraimo, David Forrest, Ian G McHale, and JD Tena. Unscripted drama: soccer audience response to suspense, surprise, and shock. Economic Inquiry, 58(2):881–896, 2020.

[6] J. L. Doob. Regularity properties of certain families of chance variables. Transactions of the American Mathematical Society, 47(3):455–486, 1940.

[7] Jeffrey Ely, Alexander Frankel, and Emir Kamenica. Suspense and surprise. Journal of Political Economy, 123(1):215–260, 2015.

[8] Daniel Kahneman, Barbara L Fredrickson, Charles A Schreiber, and Donald A Redelmeier. When more pain is preferred to less: Adding a better end. Psychological science, 4(6):401–405, 1993.

[9] Balazs Kovacs. The effect of the scoring system changes in volleyball: a model and an empirical test. Journal of Quantitative Analysis in Sports, 5(3), 2009.

[10] Daniel W Lozier. Nist digital library of mathematical functions. Annals of Mathematics and Artificial Intelligence, 38(1):105–119, 2003.

[11] Gale M. Lucas, Jonathan Gratch, Nikolaos Malandrakis, Evan Szabolowski, Eli Fessler, and Jeffrey Nichols. Goaaall!: Using sentiment in the world cup to explore theories of emotion. Image and Vision Computing, 65:58–65, 2017. ISSN 0262-8856. doi: https://doi.org/10.1016/j.imavis.2017.01.006. URL https://www.sciencedirect.com/science/article/pii/S0262885617300148 Multimodal Sentiment Analysis and Mining in the Wild Image and Vision Computing.

[12] Elchanan Mossel, Joe Neeman, and Omer Tamuz. Majority dynamics and aggregation of information in social networks. Autonomous Agents and Multi-Agent Systems, 28(3):408–429, 2014.

[13] David F Percy. A mathematical analysis of badminton scoring systems. In Operational Research Applied to Sports, pages 181–200. Springer, 2015.

[14] Phil Scarf, Rishikesh Parma, and Ian McHale. On outcome uncertainty and scoring rates in sport: The case of international rugby union. European Journal of Operational Research, 273(2):721–730, 2019.
A Proof of Lemma 4.2

Lemma 4.2 We have

\[
q_i^j = \frac{j + \alpha}{i + \alpha + \beta}
\]

\[
\Pr[S_i = j] = \frac{(\alpha + \beta - 1)(i)^{\alpha + \beta - 2}}{(\alpha + \beta + i - 1)(i + \alpha + \beta - 2)}
\]

\[
Q_i^j = \frac{2(\alpha + \beta - 1)(i)^{\alpha + \beta - 2}}{(\alpha + \beta + i)(i + \alpha + \beta - 2)}
\]

and

\[
d_i^{n-2} = \begin{cases} 
0, & j < L_n-2 \\
\frac{L_{n-1} + \alpha}{n + \alpha + \beta - 1}, & j = L_n-2 \\
\frac{1}{n + \alpha + \beta - 1}, & L_n-2 < j < U_n-2 \\
\frac{n-1-U_{n-1} + \beta}{n + \alpha + \beta - 1}, & j = U_n-2 \\
0, & j > U_n-2
\end{cases}
\]

\[
d_i^{n-1} = \begin{cases} 
0, & j < L_n-1 \\
1, & L_n-1 \leq j \leq U_n-1 \\
0, & j > U_n-1
\end{cases}
\]

Proof of Lemma 4.2 Recall that \(S_i = j\) means Alice wins \(j\) rounds in the first \(i\) rounds.

When the prior \(p\) follows \(B(\alpha, \beta)\), the posterior \(p(S_i = j)\) follows \(B(\alpha + j, \beta + i - j)\) (Claim 2.3).

We can calculate the expectation of \(p(S_{n-1} = j)\),

\[
q_i^{n-1} = \mathbb{E}[p|S_{n-1} = j] = \frac{j + \alpha}{n + \alpha + \beta - 1}
\]

We use induction to show that

\[
\forall i, j \Pr[S_i = j] = \frac{(\alpha + \beta - 1)(i)^{\alpha + \beta - 2}}{(\alpha + \beta + i - 1)(i + \alpha + \beta - 2)}
\]

For the base case \(i = 0\), note that \(j \leq i\) is also zero, the above equation holds naturally since

\[
1 = \Pr[S_0 = 0] = \frac{(\alpha + \beta - 1)(i)^{\alpha + \beta - 2}}{(\alpha + \beta + i - 1)(i + \alpha + \beta - 2)}.
\]

For the induction, we first show the relationship between the distribution of \(S_i\) and \(S_{i-1}\).

\[
\Pr[S_i = j] = \begin{cases} 
\Pr[S_{i-1} = 0] \times \frac{i + \beta - 1}{i + \alpha + \beta - 1}, & j = 0 \\
\Pr[S_{i-1} = j - 1] \times \frac{j + \beta - 1}{i + \alpha + \beta - 1} + \Pr[S_{i-1} = j] \times \frac{i-j+\beta-1}{i+\alpha+\beta-1}, & 0 < j < i \\
\Pr[S_{i-1} = i - 1] \times \frac{i+\beta-1}{i+\alpha-1}, & j = i
\end{cases}
\]

We assume that the equation holds for \(i - 1\), i.e.,

\[
\forall 0 \leq j \leq i - 1, \Pr[S_{i-1} = j] = \frac{(\alpha + \beta - 1)(i-1)^{\alpha + \beta - 2}}{(\alpha + \beta + (i-1) - 1)(i-1 + \alpha + \beta - 2)}
\]

33
Thus, the equation also holds for $i$. Therefore, we have

$$\forall i, j \Pr[S_i = j] = \frac{(\alpha + \beta - 1)\alpha_{-1}^i}{(\alpha + \beta + i - 1)\alpha_{-1}^{i+1} + (\alpha + \beta - 1)\alpha_{-1}^i}$$
Then we can calculate $Q^i_j$

$$Q^i_j = \Pr[S_i = j] \ast 2 \ast E[p|S_i = j] \ast (1 - E[p|S_i = j])$$

$$= \Pr[S_i = j] \ast 2 \ast q^i_j \ast (1 - q^i_j)$$

$$= \frac{(\alpha + \beta - 1)(j^i)_{\alpha+\beta-2}}{(\alpha + \beta + i - 1)(j^i + \alpha + \beta - 2)} \ast \frac{j + \alpha}{i + \alpha + \beta} \ast \frac{i - j + \beta}{i + \alpha + \beta}$$

$$= \frac{2(\alpha + \beta - 1)(j^i)_{\alpha+\beta-2}}{(\alpha + \beta + i)^{i+\alpha}}$$

Recall in Section 4.2, we have defined $b^{n-1}_j = \Pr[O = 1|(S_{n-1} = j)]$ and $d^{n-1}_j = \Pr[O = 1|(H_n = +) \land (S_{n-1} = j)] - \Pr[O = 1|(H_n = -) \land (S_{n-1} = j)].$

When $j < L_{n-1}$ ($j > U_{n-1}$), Alice must lose (win) regardless of the final round’s outcome. When $L_{n-1} \leq j \leq U_{n-1}$, Alice wins if and only if she wins the final round. Therefore,

$$b^{n-1}_j = \begin{cases} 0, & j < L_{n-1} \\ \frac{j + \alpha}{n + \alpha + \beta - 1}, & L_{n-1} \leq j \leq U_{n-1} \\ 1, & j > U_{n-1} \end{cases}$$

$$d^{n-1}_j = \begin{cases} 0, & j < L_{n-1} \\ \frac{L_{n-1} + \alpha}{n + \alpha + \beta - 1}, & j = L_{n-1} \\ \frac{1}{n - 1 - U_{n-1} + \beta}, & L_{n-1} < j < U_{n-1} \\ \frac{1 + \beta}{n + \alpha + \beta - 1}, & j = U_{n-1} \\ 0, & j > U_{n-1} \end{cases}$$

Moreover, recall that $d^{n-2}_j = b^{n-1}_{j+1} - b^{n-1}_j$, thus

$$d^{n-2}_j = \begin{cases} 0, & j < L_{n-2} \\ \frac{L_{n-2} + \alpha}{n + \alpha + \beta - 1}, & j = L_{n-2} \\ \frac{1}{n - 1 - U_{n-2} + \beta}, & L_{n-2} < j < U_{n-2} \\ \frac{1 + \beta}{n + \alpha + \beta - 1}, & j = U_{n-2} \\ 0, & j > U_{n-2} \end{cases}$$

\[\square\]

### B Formal Proof for Asymptotic Case

**Lemma 5.2 (Formal).** When $n$ is sufficiently large, fixed $\alpha, \beta, \mu \in (0, 1)$, we have

$$\begin{align*}
L_{n-1} &= \frac{1 - \mu}{2} \ast (1 + O(\frac{1}{n})) \\
U_{n-2} &= \frac{1 - \mu}{2} \ast (1 + O(\frac{1}{n})) \\
U_{n-1} &= \frac{1 + \mu}{2} \ast (1 + O(\frac{1}{n}))
\end{align*}$$

For all $L_{n-1} \leq j \leq U_{n-1}$, let $\theta_j := \frac{j}{n}$, we have

$$\begin{align*}
\Pr[S_{n-1} = j] &= \frac{\text{Beta}(\theta_j; \alpha, \beta)}{n} \ast (1 + O(\frac{1}{n})) \\
\Pr[S_{n-2} = j] &= \frac{\text{Beta}(\theta_j; \alpha, \beta)}{n} \ast (1 + O(\frac{1}{n})) \\
q^{n-1}_j &= \theta_j \ast (1 + O(\frac{1}{n})) \\
q^{n-2}_j &= \theta_j \ast (1 + O(\frac{1}{n}))
\end{align*}$$
and

\[ d_j^{n-2} = \begin{cases} 
0, & j < L_{n-2} \\
\frac{1-\mu}{n} * (1 + O(\frac{1}{n})), & j = L_{n-2} \\
\frac{1-\mu}{n} * (1 + O(\frac{1}{n})), & L_{n-2} < j < U_{n-2} \\
\frac{1-\mu}{n} * (1 + O(\frac{1}{n})), & j = U_{n-2} \\
0, & j > U_{n-2} 
\end{cases} \]

\[ d_j^{n-1} = \begin{cases} 
0, & j < L_{n-1} \\
1, & L_{n-1} \leq j \leq U_{n-1} \\
0, & j > U_{n-1} 
\end{cases} \]

**Proof of Lemma 5.2** First recall the definition of \( L_{n-2}, L_{n-1}, U_{n-2}, U_{n-1} \),

\[
\begin{align*}
L_{n-2} &= \frac{n-x-2}{2} \\
L_{n-1} &= \frac{n-x}{2} \\
U_{n-2} &= \frac{n+x-2}{2} \\
U_{n-1} &= \frac{n+x}{2}
\end{align*}
\]

Then by substituting \( \mu = \frac{x}{n} \),

\[
\begin{align*}
\frac{L_{n-2}}{n} &= \frac{1-\mu}{2} - \frac{1}{n} \\
\frac{L_{n-1}}{n} &= \frac{1-\mu}{2} \\
\frac{U_{n-2}}{n} &= \frac{1+\mu}{2} - \frac{1}{n} \\
\frac{U_{n-1}}{n} &= \frac{1+\mu}{2} - \frac{1}{n}
\end{align*}
\]

Given fixed \( \mu \in (0, 1) \),

\[
\begin{align*}
\frac{L_{n-2}}{n} &= \frac{1-\mu}{2} + (1 - \frac{1}{n}) / \frac{1-\mu}{2} = \frac{1-\mu}{2} + (1 + O(\frac{1}{n})) \\
\frac{L_{n-1}}{n} &= \frac{1-\mu}{2} \\
\frac{U_{n-2}}{n} &= \frac{1+\mu}{2} + (1 - \frac{1}{n}) / \frac{1+\mu}{2} = \frac{1+\mu}{2} + (1 + O(\frac{1}{n})) \\
\frac{U_{n-1}}{n} &= \frac{1+\mu}{2} + (1 - \frac{1}{n}) / \frac{1+\mu}{2} = \frac{1+\mu}{2} + (1 + O(\frac{1}{n}))
\end{align*}
\]

Second, recall the formula of \( \Pr[S_{n-1} = j] \) and \( \Pr[S_{n-2} = j] \).

\[
\begin{align*}
\Pr[S_{n-1} = j] &= \frac{(\alpha+\beta-1)\binom{n-1}{j}^2}{(\alpha+\beta+n-2)\binom{n-1}{j}^2} \\
\Pr[S_{n-2} = j] &= \frac{(\alpha+\beta-1)\binom{n-2}{j}^2}{(\alpha+\beta+n-3)\binom{n-2}{j}^2}
\end{align*}
\]

To show that \( \Pr[S_{n-1} = j] = \frac{\beta^{Beta(\theta;\alpha,\beta)}}{n} \), we calculate the ratio between \( \Pr[S_{n-1} = j] \) and \( \frac{\beta^{Beta(\theta;\alpha,\beta)}}{n} \).
Proof of Claim B.1. Since \( \frac{n}{j} \leq \frac{n}{20} \), then we need to prove that each item of the above formula is \( \Theta(n) \). For this, we use Gautschi’s inequality (see 5.6.4 in Lozier [10]).

Then we need to prove that each item of the above formula is \( 1 + O\left(\frac{1}{n}\right) \).

Since \( L_{n-1} \leq j \leq U_{n-1} \text{ and } L_{n-1} = \frac{1}{n} \), \( U_{n-1} = \frac{1}{n} \), we have \( j = \Theta(n) \) and \( n - j = \Theta(n) \). Therefore, \( \frac{j + 1}{j} = 1 + O\left(\frac{1}{n}\right) \).

It’s left to prove that \( \frac{j+1}{j} \), \( \frac{n-j}{n-j} \), and \( \frac{n-alpha-beta}{n-alpha-beta} \) are all \( 1 + O\left(\frac{1}{n}\right) \). We give the following claim:

Claim B.1. For any \( x = \Theta(n), t = \Theta(1), \frac{x^t}{t^x} = 1 + O\left(\frac{1}{n}\right) \)

Proof of Claim B.1. When \( t \) is an integer, we have

\[
x^t \leq x^t = \frac{\Gamma(x+t)}{\Gamma(x)} = \prod_{i=1}^{t}(x+i-1) \leq (x+t-1)^t,\]

then

\[
1 \leq \frac{x^t}{x^t} \leq \left(1 + \frac{t-1}{x}\right)^t.
\]

Since \( 1 + \frac{t-1}{x} = 1 + O\left(\frac{1}{n}\right) \), then \( (1 + \frac{t-1}{x})^t = (1 + O\left(\frac{1}{n}\right))^t = 1 + O\left(\frac{1}{n}\right) \), and we have \( \frac{x^t}{x^t} = 1 + O\left(\frac{1}{n}\right) \).

When \( t \) is not an integer, we need to use Gautschi’s inequality (see 5.6.4 in Lozier [10])

\[
\forall s \in (0, 1), x^{1-s} < \frac{\Gamma(x+1)}{\Gamma(x+s)} < (x+1)^{1-s}.
\]

\(^{21}\) We define \( x^t \) as the continuous generalization of rising factorials, i.e. \( x^t = \frac{\Gamma(x+t)}{\Gamma(x)} \).
Let $c$ be the smallest integer larger than $t$, then we have
\[ x^c < x^\bar{t} = \frac{\Gamma(x + c)}{\Gamma(x)} < (x + c - 1)^c \quad \text{(Due to previous analysis)} \]
\[ (x + c - 1)^{c-t} < \frac{\Gamma(x + c)}{\Gamma(x + t)} < (x + c)^{c-t} \quad \text{(Gautschi's inequality)} \]
\[ \Rightarrow \quad \frac{x^c}{(x + c)^{c-t}} < \frac{x^\bar{t}}{\Gamma(x + t)} < (x + c - 1)^t \]

then
\[ \frac{x}{x + c} < \left( \frac{x}{x + c} \right)^{c-t} < \frac{x^\bar{t}}{x^t} < (1 + \frac{c-1}{x})^t < (1 + \frac{c-1}{x})^c \]

Since $\frac{x}{x+c} = 1 + O(\frac{1}{n})$ and $(1 + \frac{c-1}{x}) = 1 + O(\frac{1}{n})$, then $(1 + \frac{c-1}{x})^c = 1 + O(\frac{1}{n})$.

Thus we proved $\frac{x^\bar{t}}{x^t} = 1 + O(\frac{1}{n})$. □

Based on Claim B.1 we have \( \frac{(j^\frac{n-1}{j-1})}{(j-1)^*} = 1 + O(\frac{1}{n}) \), \( \frac{(n-j)^{\frac{n-1}{n-j}}}{(n-j)^*} = 1 + O(\frac{1}{n}) \) and \( \frac{n}{n+\beta-1} = \frac{1}{1+O(\frac{1}{n})} = 1 + O(\frac{1}{n}) \)

Then
\[ \Pr[S_{n-1} = j] = \frac{\text{Beta}(\theta_j;\alpha,\beta)}{n} \]
\[ = (1 + O(\frac{1}{n})) \times (1 + O(\frac{1}{n})) \times (1 + O(\frac{1}{n})) \times (1 + O(\frac{1}{n})) \times (1 + O(\frac{1}{n})) \]
\[ = 1 + O(\frac{1}{n}) \]

In order to prove \( \Pr[S_{n-2} = j] = \frac{\text{Beta}(\theta_j;\alpha,\beta)}{n} \times (1 + O(\frac{1}{n})) \), we first prove \( \Pr[S_{n-2} = j] = \Pr[S_{n-1} = j] \times (1 + O(\frac{1}{n})) \)

\[ \Pr[S_{n-2} = j] / \Pr[S_{n-1} = j] = \frac{(\alpha + \beta + n - 3)(n+\alpha+\beta-3)(j-1)}{(\alpha + \beta + n - 2)(n+\alpha+\beta-1)(j-1)} \]
\[ = \frac{\alpha + \beta + n - 3}{\alpha + \beta + n - 2} \times \frac{n + \alpha + \beta - 3}{n - j - \beta - 2} \times \frac{n - j - 1}{n - 1} \]
\[ = \frac{\alpha + \beta + n - 3}{\alpha + \beta + n - 2} \times \frac{n + \alpha + \beta - 3}{n - j - \beta - 2} \times \frac{n - j - 1}{n - 1} \]
\[ = (1 + O(\frac{1}{n})) \times (1 + O(\frac{1}{n})) \times (1 + O(\frac{1}{n})) \quad (\alpha, \beta \text{ are constants}) \]
\[ = 1 + O(\frac{1}{n}) \]

Since \( \Pr[S_{n-1} = j] = \frac{\text{Beta}(\theta_j;\alpha,\beta)}{n} \times (1 + O(\frac{1}{n})) \), we proved \( \Pr[S_{n-2} = j] = \frac{\text{Beta}(\theta_j;\alpha,\beta)}{n} \times (1 + O(\frac{1}{n})) \)

Third, recall the formula of \( \phi_j^{n-1} \) and \( \phi_j^{n-2} \)

\[
\begin{align*}
\phi_j^{n-1} &= \frac{j+\alpha}{n+\alpha+\beta-1} \\
\phi_j^{n-2} &= \frac{j+\alpha}{n+\alpha+\beta-2}
\end{align*}
\]
We have
\[
\begin{aligned}
q_j^{n-1}/\theta_j &= \frac{j+\alpha}{n+\alpha+\beta-1} / \frac{j}{n} = \frac{j+\alpha}{j} \cdot \frac{n}{n+\alpha+\beta-1} \\
q_j^{n-2}/\theta_j &= \frac{j+\alpha}{n+\alpha+\beta-2} / \frac{j}{n} = \frac{j+\alpha}{j} \cdot \frac{n}{n+\alpha+\beta-2}
\end{aligned}
\]

Thus
\[
\begin{aligned}
q_j^{n-1}/\theta_j &= (1 + O(\frac{1}{n})) \cdot (1 + O(\frac{1}{n})) = 1 + O(\frac{1}{n}) \\
q_j^{n-2}/\theta_j &= (1 + O(\frac{1}{n})) \cdot (1 + O(\frac{1}{n})) = 1 + O(\frac{1}{n})
\end{aligned}
\]
\[
\Rightarrow \begin{cases}
q_j^{n-1}/\theta_j = (1 + O(\frac{1}{n})) \\
q_j^{n-2}/\theta_j = (1 + O(\frac{1}{n}))
\end{cases}
\]

Finally, recall the formula of \(d_j^{n-2}\)
\[
d_j^{n-2} = \begin{cases}
0, & j < L_{n-2} \\
\frac{L_{n-1}+\alpha}{n+\alpha+\beta-1}, & j = L_{n-2} \\
\frac{1}{n+\alpha+\beta-1}, & L_{n-2} < j < U_{n-2} \\
\frac{n-1-U_{n-1}+\beta}{n+\alpha+\beta-1}, & j = U_{n-2} \\
0, & j > U_{n-2}
\end{cases}
\]

We have
\[
d_j^{n-2} = \begin{cases}
\frac{L_{n-1}+\alpha}{n+\alpha+\beta-1} = \frac{L_{n-1}}{n} \cdot (1 + O(\frac{1}{n})), & j = L_{n-2} \\
\frac{1}{n+\alpha+\beta-1} = \frac{1}{n} \cdot (1 + O(\frac{1}{n})), & L_{n-2} < j < U_{n-2} \\
\frac{n-1-U_{n-1}+\beta}{n+\alpha+\beta-1} = \frac{n-U_{n-1}}{n} \cdot (1 + O(\frac{1}{n})), & j = U_{n-2} \\
0, & j > U_{n-2}
\end{cases}
\]

By substituting the value of \(L_{n-2}\) and \(U_{n-2}\), we get
\[
d_j^{n-2} = \begin{cases}
0, & j < L_{n-2} \\
\frac{1}{n} \cdot (1 + O(\frac{1}{n})), & j = L_{n-2} \\
\frac{1}{n} \cdot (1 + O(\frac{1}{n})), & L_{n-2} < j < U_{n-2} \\
\frac{1}{n} \cdot (1 + O(\frac{1}{n})), & j = U_{n-2} \\
0, & j > U_{n-2}
\end{cases}
\]

We use Lemma 5.2 to prove
\[
\mathbb{E}[\Delta_B(n \ast \mu)] = Z_{\alpha, \beta, n}(\mu) \ast (1 + O(\frac{1}{n})).
\]

We prove the four parts separately.

**Penultimate Round (at point \(L_{n-2}\))** We need to prove
\[
\Pr[S_{n-2} = L_{n-2}] \ast 2q_{L_{n-2}}^{n-2} \ast (1-q_{L_{n-2}}^{n-2}) \ast d_{L_{n-2}}^{n-2} = \frac{\text{Beta}(\frac{1+\mu}{2}; \alpha, \beta) \ast 2 \ast \frac{1+\mu}{2} \ast \frac{1-\mu}{2} \ast \frac{1-\mu}{2} \ast (1 + O(\frac{1}{n}))}{n}
\]
Recall Lemma 5.2 we have
\[
\begin{aligned}
\Pr[S_{n-2} = L_{n-2}] &= \text{Beta}(\frac{1+\mu}{2}; \alpha, \beta) \\
q_{L_{n-2}}^{n-2} &= \frac{1+\mu}{2} = (1 + O(\frac{1}{n})) \\
1 - q_{L_{n-2}}^{n-2} &= \frac{1+\mu}{2} = (1 + O(\frac{1}{n})) \\
d_{L_{n-2}}^{n-2} &= \frac{1+\mu}{2} = (1 + O(\frac{1}{n}))
\end{aligned}
\]

39
By substituting the above items into the **penultimate round (at point \( L_{n-2} \))** part, we get

\[
\Pr[S_{n-2} = L_{n-2}] * 2q_{L_{n-2}}^{n-2} * (1 - q_{L_{n-2}}^{n-2}) * d_{L_{n-2}}^{n-2} = \frac{\text{Beta}(\frac{1 + \mu}{2}; \alpha, \beta)}{n} \ast (1 + O(\frac{1}{n})) \ast 2 \ast \frac{1 - \mu}{2} \ast (1 + O(\frac{1}{n})) \ast \frac{1 + \mu}{2} \ast (1 + O(\frac{1}{n})) \ast \frac{1 - \mu}{2} \ast (1 + O(\frac{1}{n}))
\]

**Penultimate Round (at point \( U_{n-2} \))** We need to prove

\[
\Pr[S_{n-2} = U_{n-2}] * 2q_{U_{n-2}}^{n-2} * (1 - q_{U_{n-2}}^{n-2}) * d_{U_{n-2}}^{n-2} = \frac{\text{Beta}(\frac{1 + \mu}{2}; \alpha, \beta)}{n} \ast 2 \ast \frac{1 + \mu}{2} \ast \frac{1 - \mu}{2} \ast \frac{1 - \mu}{2} \ast (1 + O(\frac{1}{n}))
\]

Recall Lemma 5.2 we have

\[
\begin{align*}
\Pr[S_{n-2} = U_{n-2}] & = \frac{\text{Beta}(\frac{1 + \mu}{2}; \alpha, \beta)}{n} \ast (1 + O(\frac{1}{n})) \\
q_{U_{n-2}}^{n-2} & = \frac{1 + \mu}{2} \ast (1 + O(\frac{1}{n})) \\
1 - q_{U_{n-2}}^{n-2} & = \frac{1 - \mu}{2} \ast (1 + O(\frac{1}{n})) \\
d_{U_{n-2}}^{n-2} & = \frac{1 - \mu}{2} \ast (1 + O(\frac{1}{n}))
\end{align*}
\]

By substituting the above items into the **penultimate round (at point \( U_{n-2} \))** part, we get

\[
\Pr[S_{n-2} = U_{n-2}] * 2q_{U_{n-2}}^{n-2} * (1 - q_{U_{n-2}}^{n-2}) * d_{U_{n-2}}^{n-2} = \frac{\text{Beta}(\frac{1 + \mu}{2}; \alpha, \beta)}{n} \ast (1 + O(\frac{1}{n})) \ast 2 \ast \frac{1 + \mu}{2} \ast \frac{1 - \mu}{2} \ast \frac{1 - \mu}{2} \ast (1 + O(\frac{1}{n}))
\]

To prove the remaining parts, we need to extend our results for \( \theta_j \) to \( \theta \in [\theta_j, \theta_{j+1}] \). Note that \( \forall \theta \in [\theta_j, \theta_{j+1}], \frac{L_{n-1}}{n} < \theta < \frac{U_{n-1}}{n} \), we have

\[
\begin{align*}
\text{Beta}(\theta_j; \alpha, \beta) & = \text{Beta}(\theta; \alpha, \beta) * O(1 + \frac{1}{n})^{19} \\
q_{\theta_j, n}^{n-2} & = q_{\theta, n}^{n-2} * O(1 + \frac{1}{n}) \\
q_{\theta_j, n}^{n-1} & = q_{\theta, n}^{n-1} * O(1 + \frac{1}{n}) \\
d_{\theta_j, n}^{n-2} & = d_{\theta, n}^{n-2} * O(1 + \frac{1}{n}) \\
d_{\theta_j, n}^{n-1} & = d_{\theta, n}^{n-1} * O(1 + \frac{1}{n})
\end{align*}
\]

We are ready to prove the last two parts.

\(^{22}\text{Note that the derivative of } \text{Beta}(\theta; \alpha, \beta) \text{ is bounded.}\)
Penultimate Round (between $L_{n-2}$ and $U_{n-2}$)  

Here we need to prove

$$
\sum_{j=L_{n-2}+1}^{U_{n-2}-1} \Pr[S_{n-2} = j] \times 2q_j^n \times (1 - q_j^n) \times d_j^n = \int_{\frac{1}{U_{n-2}}}^{\frac{1}{L_{n-2}}} \text{Beta}(\theta; \alpha, \beta) \times 2 \times (1 - \theta) \times \frac{1}{n} \times d\theta \times (1 + O(\frac{1}{n}))
$$

Recall Lemma 5.2 we have

$$
\begin{align*}
\Pr[S_{n-2} = j] &= \frac{\text{Beta}(\theta_j; \alpha, \beta)}{n} \times (1 + O(\frac{1}{n})) \\
q_j^n &= \theta_j \times (1 + O(\frac{1}{n})) \\
1 - q_j^n &= (1 - \theta_j) \times (1 + O(\frac{1}{n})) \\
d_j^n &= \frac{1}{n} \times (1 + O(\frac{1}{n}))
\end{align*}
$$

By substituting the above items into penultimate round (between $L_{n-2}$ and $U_{n-2}$) part, we get

$$
\sum_{j=L_{n-2}+1}^{U_{n-2}-1} \Pr[S_{n-2} = j] \times 2q_j^n \times (1 - q_j^n) \times d_j^n = \\
\sum_{j=L_{n-2}+1}^{U_{n-2}-1} \frac{\text{Beta}(\theta_j; \alpha, \beta)}{n} \times (1 + O(\frac{1}{n})) \times 2 \times \theta_j \times (1 + O(\frac{1}{n})) \times (1 - \theta_j) \times (1 + O(\frac{1}{n})) \times \frac{1}{n} \times (1 + O(\frac{1}{n}))
$$

Then we have

$$
\sum_{j=L_{n-2}+1}^{U_{n-2}-1} \frac{\text{Beta}(\theta_j; \alpha, \beta)}{n} \times 2 \times \theta_j \times (1 - \theta_j) \times \frac{1}{n} = \\
\sum_{j=L_{n-2}+1}^{U_{n-2}-1} \frac{\text{Beta}(\theta_j; \alpha, \beta)}{n} \times 2 \times \theta_j \times (1 - \theta_j) \times \frac{1}{n} \times d\theta \times (1 + O(\frac{1}{n}))
$$

Notes that

$$
\frac{\text{Beta}(\theta_j; \alpha, \beta)}{n} \times 2 \times \theta_j \times (1 - \theta_j) \times \frac{1}{n} = \\
(\theta_{j+1} - \theta_j) \times (\text{Beta}(\theta_j; \alpha, \beta) \times 2 \times \theta_j \times (1 - \theta_j)) \times \frac{1}{n}
$$

$$
= \int_{\theta_j}^{\theta_{j+1}} (\text{Beta}(\theta_j; \alpha, \beta) \times 2 \times \theta_j \times (1 - \theta_j)) \times \frac{1}{n} \times d\theta
$$

$$
= \int_{\theta_j}^{\theta_{j+1}} (\text{Beta}(\theta; \alpha, \beta) \times (1 + O(\frac{1}{n})) \times 2 \times \theta \times (1 + O(\frac{1}{n})) \times (1 - \theta) \times (1 + O(\frac{1}{n})) \times \frac{1}{n}) \times d\theta
$$

$$
= \int_{\theta_j}^{\theta_{j+1}} (\text{Beta}(\theta; \alpha, \beta) \times 2 \times \theta \times (1 - \theta) \times \frac{1}{n}) \times d\theta \times (1 + O(\frac{1}{n}))
$$

Finally, we have

$$
\left[13\right] = \sum_{j=L_{n-2}+1}^{U_{n-2}-1} \int_{\theta_j}^{\theta_{j+1}} (\text{Beta}(\theta; \alpha, \beta) \times 2 \times \theta \times (1 - \theta) \times \frac{1}{n}) \times d\theta \times (1 + O(\frac{1}{n}))
$$

$$
= \int_{\frac{1}{U_{n-2}}}^{\frac{1}{L_{n-2}}} (\text{Beta}(\theta; \alpha, \beta) \times 2 \times \theta \times (1 - \theta) \times \frac{1}{n}) \times d\theta \times (1 + O(\frac{1}{n}))
$$

$$
= \frac{1}{n} \left( \int_{\frac{1}{U_{n-2}}}^{\frac{1}{L_{n-2}}} \text{Beta}(\theta; \alpha, \beta) \times 2 \times \theta \times (1 - \theta) \times d\theta \times (1 + O(\frac{1}{n})) \right)
$$

$$
= \int_{\frac{1}{U_{n-2}}}^{\frac{1}{L_{n-2}}} (\text{Beta}(\theta; \alpha, \beta) \times 2 \times \theta \times (1 - \theta) \times \frac{1}{n}) \times d\theta \times (1 + O(\frac{1}{n}))
$$

\[41\]
Final Round  Finally, we need to prove
\[ \sum_{j=L_{n-1}}^{U_{n-1}} \Pr[S_{n-1} = j] \cdot 2q_j^{n-1} \cdot (1-q_j^{n-1}) \cdot d_j^{n-1} = \int_{1-\mu}^{1+\mu} \text{Beta}(\theta; \alpha, \beta) \cdot 2 \cdot \theta \cdot (1-\theta) \, d\theta \cdot (1 + O(\frac{1}{n})) \]

The proof is similar to the penultimate round (between \(L_{n-2}\) and \(U_{n-2}\)) part. Recall Lemma 5.2, we have
\[
\begin{align*}
\Pr[S_{n-1} = j] &= \frac{\text{Beta}(\theta; \alpha, \beta)}{n} \cdot (1 + O(\frac{1}{n})) \\
q_j^{n-1} &= \theta \cdot (1 + O(\frac{1}{n})) \\
1 - q_j^{n-1} &= (1 - \theta) \cdot (1 + O(\frac{1}{n})) \\
d_j^{n-1} &= 1
\end{align*}
\]

By substituting the above items into final round part, we get
\[
\sum_{j=L_{n-1}}^{U_{n-1}} \Pr[S_{n-1} = j] \cdot 2q_j^{n-1} \cdot (1-q_j^{n-1}) \cdot d_j^{n-1}
= \sum_{j=L_{n-1}}^{U_{n-1}} \frac{\text{Beta}(\theta; \alpha, \beta)}{n} \cdot (1 + O(\frac{1}{n})) \cdot 2 \cdot \theta \cdot (1 + O(\frac{1}{n})) \cdot (1 - \theta) \cdot (1 + O(\frac{1}{n})) \cdot \frac{1}{n} \cdot (1 + O(\frac{1}{n}))
= \sum_{j=L_{n-1}}^{U_{n-1}} \frac{\text{Beta}(\theta; \alpha, \beta)}{n} \cdot 2 \cdot \theta \cdot (1 - \theta) \cdot \frac{1}{n} \cdot (1 + O(\frac{1}{n}))
= \sum_{j=L_{n-1}}^{U_{n-1}} \int_{\theta_j}^{\theta_{j+1}} (\text{Beta}(\theta; \alpha, \beta) \cdot 2 \cdot \theta \cdot (1 - \theta) \cdot \frac{1}{n}) \, d\theta \cdot (1 + O(\frac{1}{n}))
= \int_{\frac{U_{n-1} - 1}{n}}^{\frac{L_{n-1} + 1}{n}} (\text{Beta}(\theta; \alpha, \beta) \cdot 2 \cdot \theta \cdot (1 - \theta) \cdot \frac{1}{n}) \, d\theta \cdot (1 + O(\frac{1}{n}))
\text{(Recall formula (14))}
= \frac{1}{n} \left( \int_{\frac{1-\mu}{n}}^{\frac{1+\mu}{n}} (\text{Beta}(\theta; \alpha, \beta) \cdot 2 \cdot \theta \cdot (1 - \theta) \cdot \frac{1}{n}) \, d\theta + O(\frac{1}{n}) \right) \cdot (1 + O(\frac{1}{n}))
(Beta(\theta; \alpha, \beta) \cdot 2 \cdot \theta \cdot (1 - \theta) \text{ is bounded by two positive constants})
= \int_{\frac{1-\mu}{n}}^{\frac{1+\mu}{n}} (\text{Beta}(\theta; \alpha, \beta) \cdot 2 \cdot \theta \cdot (1 - \theta) \cdot \frac{1}{n}) \, d\theta \cdot (1 + O(\frac{1}{n}))
\]

C Property of \(F(x)\) and \(G(\mu)\)

Lemma 4.6  When \(p \geq \frac{1}{2}\), \(F(x) = 0\) has a trivial solution at \(x = 1\), and has a non-trivial solution \(\hat{x} \in (1, 2np - n + 1)\) if and only if \(p > \frac{1}{2}\) and \(n > \frac{1}{(\frac{1}{2} - p) \ln(\frac{1}{2p})}\). There is no other solution. Moreover, when \(x \in (1, \hat{x})\), \(F(x) > 0\), when \(\hat{x} < n - 1\) and \(x \in (\hat{x}, n - 1)\), \(F(x) < 0\).

Besides, let \(a = 2np - n - 2\), when \(p > \frac{1}{1 + (a + 1)^{1/3}}\), the non-trivial solution of \(F(x) = 0\) is in \((2np - n - 1, 2np - n + 1)\).

Proof of Lemma 4.6  We first define two shorthand notations:
\[
\begin{align*}
b &= 2np - n \\
t &= x - 1
\end{align*}
\]
We have $F(t+1) = (b-t)p^t + (-b-t)(1-p)^t$, where $t \in [0, n)$. When $p = \frac{1}{2}$, $F(t+1) = -2t(\frac{1}{2})^t$ thus only has a trivial solution at $t = 0$. We only need to consider the case of $p > \frac{1}{2}$. Here $b > 0$. $F(t+1)|_{t=0} = 0$, and $F(t+1)|_{t=\alpha} < 0$ since $(b-t)p^t \leq 0$ and $(-b-t)(1-p)^t |_{t=\alpha} < 0$. Then to analyze $F(t+1) = 0$'s other solutions, we only need to discuss $t \in [0, b)$, i.e. $x \in [1, 2np-n+1]$.

$$F(x) \geq 0$$

$$\iff \frac{b-t}{b+t} \geq \left(\frac{1-p}{p}\right)^t$$

$$\iff \ln \frac{b-t}{b+t} \geq t \ln \frac{1-p}{p}$$

Let $f(t) = \ln \frac{b-t}{b+t} - t \ln \frac{1-p}{p}$, then we have $F(x) > 0 \iff f(t) > 0$, and $F(x) = 0 \iff f(t) = 0$.

In addition to the trivial solution $f(0) = 0$, to analyze other solutions in $t \in (0, b)$, we calculate the first and second derivative of $f(t)$,

$$\frac{d}{dt} f(t) = -\frac{2b}{b^2 - t^2} - \ln \frac{1-p}{p}$$

$$\frac{d^2}{dt^2} f(t) = -\frac{4bt}{(b-t)^2(b+t)^2}$$

Recall that $b = 2np - n = (2p-1)n > 0$ since $n > 1$, $p > \frac{1}{2}$. Thus, $\frac{d^2}{dt^2} f(t) < 0$ when $t \in (0, b)$, i.e. $f(t)$ is concave. Moreover, $f(0) = 0$ and $\lim_{t \to b^+} f(t) = -\infty$. Therefore, $f(t) = 0$ has a non-trivial solution $\tilde{t} \in (0, b)$ if and only if $\frac{d}{dt} f(0) = -\frac{2}{2np-n} - \ln \frac{1-p}{p} > 0$, i.e. $n > \frac{1}{(\frac{1}{2}-p)\ln(\frac{1}{2p})}$.

Moreover, $f(t) > 0$ when $t < \tilde{t}$ and $f(t) < 0$ when $t > \tilde{t}$.

This implies that when $n > \frac{1}{(\frac{1}{2}-p)\ln(\frac{1}{2p})}$, $F(x)$ has a non-trivial solution, denoted by $\tilde{x}$ and $F(x) > 0$ when $1 < x < \tilde{x}$ and $F(x) < 0$ when $\tilde{x} < x < 2np - n + 1$.

It’s left to show that when $p > \frac{1}{1+(a+1)^{-\frac{1}{2}}}$, where $a = 2np - n - 2$, the non-trivial solution of $F(x) = 0$ is in $(2np-n-1, 2np-n+1)$. Note that $F(2np-n+1) = 2(n-2np)(1-p)^{2np-n} < 0$. Thus, when $F(2np-n-1) > 0$, i.e.,

$$2p^{2np-n-2} + 2(n-2np-1)(1-p)^{2np-n-2} > 0,$$

there is a solution $\tilde{x}$ of $F(x) = 0$ in $(2np-n-1, 2np-n+1)$. The above inequality is equivalent to $p > \frac{1}{1+(a+1)^{-\frac{1}{2}}}$.

Thus, when $p > \frac{1}{1+(a+1)^{-\frac{1}{2}}}$, $F(x) = 0$ has a solution $\tilde{x} \in (2np-n-1, 2np-n+1)$, which means the difference between $2np - n$ ("expected lead") and $\tilde{x}$ (the solution of $F(x) = 0$) is less than 1, i.e. $|\tilde{x} - (2np - n)| < 1$.

\[\square\]

**Lemma 5.3** (Property of $G(\mu)$). For all $\alpha \geq \beta$, when $n$ is sufficiently large, $G(0) > 0$ and $G(1) < 0$. $G(\mu) = 0, \mu \in [0, 1]$ has a unique solution and the solution is in $(0, \frac{\alpha - \beta}{\alpha + \beta}\frac{1}{\mu})$.
Moreover, for all $0 < \epsilon < \frac{(\alpha - \beta)H+1}{(\alpha + \beta)\mathbb{H}-1}$, when $\alpha - \beta > \frac{\log(\frac{2(\alpha - \beta)H+1}{(\alpha + \beta)\mathbb{H}-1}-\epsilon)-\log \epsilon}{\log(1+\frac{(\alpha - \beta)H+1}{(\alpha + \beta)\mathbb{H}-1}-\epsilon)-\log(1-(\frac{(\alpha - \beta)H+1}{(\alpha + \beta)\mathbb{H}-1}-\epsilon))}$, the solution is within $\left(\frac{(\alpha - \beta)\mathbb{H}+1}{(\alpha + \beta)\mathbb{H}-1} - \epsilon, \frac{(\alpha - \beta)\mathbb{H}+1}{(\alpha + \beta)\mathbb{H}-1}\right) \approx \left(\frac{\alpha - \beta}{\alpha + \beta} - \epsilon, \frac{\alpha - \beta}{\alpha + \beta}\right)$.

**Proof of Lemma 5.3.** We first define three shorthand notations:

\[
\begin{aligned}
w &:= \alpha - \beta \geq 0 \\
\ell &:= \frac{(\alpha - \beta)H+1}{(\alpha + \beta)\mathbb{H}-1} \\
f &:= -\frac{(\alpha + \beta)\mathbb{H}+1}{(\alpha + \beta)\mathbb{H}-1}
\end{aligned}
\]

Notice that

\[
\ell - f = \frac{2}{(\alpha + \beta)\mathbb{H}-1}.
\]

In the asymptotic case, $\mathbb{H}$ can be sufficiently large, thus $\ell \approx f, \ell - f > 0$ and both $\ell, f \in [0, 1]$ when $n$ is sufficiently large.

We have

\[
G(\mu) = (1 + \mu)^w (\ell - \mu) + (1 - \mu)^w (-f - \mu)
\]

First we notice that $G(0) = \ell - f > 0$ and $G(1) = 2^w (\ell - 1) < 0$.

\[
G(\mu) = 0 \iff \frac{(1 + \mu)^w}{(1 - \mu)^w} = \frac{f + \mu}{\ell - \mu}
\]

Thus $G(\mu) = 0$’s solution must be less than $\ell$. We only need to consider $\mu \in (0, \ell)$. Then we can use logarithm:

\[
G(\mu) = 0 \iff w \ln \frac{1 + \mu}{1 - \mu} = \ln \frac{f + \mu}{\ell - \mu}
\]

Let $g(\mu) := w \ln \frac{1 + \mu}{1 - \mu} - \ln \frac{f + \mu}{\ell - \mu}$. Note that $g(0) = \ln \ell - \ln f > 0$ and $\lim_{\mu \to -\epsilon} g(\mu) = -\infty$. Thus, both $g(\mu) = 0$ and $G(\mu) = 0$ must have a solution in $(0, \ell)$. It’s left to show that this solution is unique.

\[
\frac{d}{d\mu} g(\mu) = w \frac{2}{1 - \mu^2} - \frac{\ell + f}{(\ell - \mu)(f + \mu)}
\]

\[
\frac{d}{d\mu} g(\mu) = 0 \iff 2w(\ell - \mu)(f + \mu) = (\ell + f)(1 - \mu^2)
\]

Let

\[
\phi(\mu) := 2w(\ell - \mu)(f + \mu) - (\ell + f)(1 - \mu^2) = (\ell + f - 2w)\mu^2 + 2w(\ell - f)\mu + 2\ell w f - \ell - f.
\]

We will show that at most one of $\phi(\mu) = 0$’s solutions is $\geq 0$. Intuitively, $\phi(\mu)$ is symmetric regarding $\mu = 0$ approximately when $n$ is sufficiently large. Thus, one of the solutions must be less than 0. Formally, note that $\phi(\mu) = 0$’s solutions are

\[
-b \pm \sqrt{b^2 - 4ac} \\
\frac{2a}{2a}
\]

where $a = \ell + f - 2w = \frac{2(\alpha - \beta)\mathbb{H}}{(\alpha + \beta)\mathbb{H}-1} - 2(\alpha - \beta) \leq 0$ since $\alpha, \beta \geq 1$. 

44
If $a = 0$, i.e., $\alpha = \beta$, $\phi(\mu) = 0$ only has one solution.
If $a < 0$, i.e., $\alpha > \beta$, note that
\[
c = 2w\ell f - \ell - f = 2(\alpha - \beta) \frac{H + 1 + (\alpha - \beta) - 1}{(\alpha + \beta) H - 1} - \frac{H + 1 + (\alpha - \beta) - 1}{(\alpha + \beta) H - 1} = \frac{2(\alpha - \beta)}{(\alpha + \beta) H - 1} \left( ((\alpha - \beta) H + 1)((\alpha - \beta) H - 1) - H((\alpha + \beta) H - 1) \right) = \frac{2(\alpha - \beta)}{(\alpha + \beta) H - 1} \left( H^2((\alpha - \beta)^2 - (\alpha + \beta)) + H - 1 \right)
\]
If $(\alpha - \beta)^2 \geq (\alpha + \beta)$, when $n$ is sufficiently large, $c > 0$, recall that $a < 0$, then $-b + \sqrt{b^2 - 4ac} > 0$, one solution $\frac{-b + \sqrt{b^2 - 4ac}}{2a} < 0$.

If $(\alpha - \beta)^2 < (\alpha + \beta)$, $c$ goes to a negative value when $n$ goes to infinity while $b$ goes to 0. In this case, $b^2 - 4ac$ goes to a negative value when $n$ goes to infinity. Thus, when $n$ is sufficiently large, $\phi(\mu) = 0$ does not have any solution.

Therefore, at most one of $\phi(\mu) = 0$’s solutions is in $[0, 1]$. This shows that at most one of $\frac{d}{d\mu} g(\mu) = 0$’s solutions is in $[0, 1]$ since $\frac{d}{d\mu} g(\mu) = 0$ is equivalent to $\phi(\mu) = 0$.
\[
\frac{d}{d\mu} g(\mu)|_{\mu=0} = 2w - \frac{\ell + f}{\ell f}
\]
\[
\lim_{\mu \to \ell^{-}} \frac{d}{d\mu} g(\mu) = -\infty
\]
Thus, $\frac{d}{d\mu} g(\mu)$ is either always negative or first positive and then negative. This shows that $g(\mu)$ is either always decreasing or first increasing and then decreasing. Recall that $g(0) = \ln \ell - \ln f > 0$ and $\lim_{\mu \to \ell^{-}} g(\mu) = -\infty$. Therefore, $g(\mu) = 0$ must have a unique solution in $(0, \ell)$.

Moreover, for all $0 < \epsilon < \ell$ if $w$ is sufficiently large such that $g(\ell - \epsilon) > 0$, the unique solution is in $(\ell - \epsilon, \ell)$. By algebraic computations, we have $g(\ell - \epsilon) > 0$ is equivalent to
\[
\alpha - \beta > \frac{\ln(\frac{(2(\alpha - \beta)) H - 1}{(\alpha + \beta) H - 1}) - \ln \epsilon}{\ln(1 + \frac{(\alpha - \beta) H - 1}{(\alpha + \beta) H - 1}) - \ln(1 - \frac{(\alpha - \beta) H - 1}{(\alpha + \beta) H - 1} - \epsilon))}.
\]
Thus, the results follow.

\[\square\]

**D Additional Proofs**

**Claim 4.3** (Local Maximum → Optimal Bonus). *If there exists $\tilde{x} \in (0, n + 1)$ such that for all $1 \leq x < \tilde{x}$, $E[\Delta_{B}(x + 1)] \geq E[\Delta_{B}(x - 1)]$ and when $\tilde{x} \leq n - 1$, for all $\tilde{x} \leq x \leq n - 1$, $E[\Delta_{B}(x + 1)] \leq E[\Delta_{B}(x - 1)]$, then $RD(\tilde{x})$ is the optimal bonus.*

**Proof of Claim 4.3** We will prove that $E[\Delta_{B}(RD(\tilde{x}))]$ is the maximum, i.e., for all $x \in \mathcal{X}(n)$, $E[\Delta_{B}(RD(\tilde{x}))] \leq E[\Delta_{B}(RD(\tilde{x}))]$.

First, we prove that 1) when $RD(\tilde{x}) - 2 \geq 0$, $E[\Delta_{B}(RD(\tilde{x}))]$ is larger than or equal to $E[\Delta_{B}(RD(\tilde{x}) - 2)]$; 2) when $RD(\tilde{x}) + 2 \leq n$, $E[\Delta_{B}(RD(\tilde{x}))]$ is larger than or equal to $E[\Delta_{B}(RD(\tilde{x}) + 2)]$.
Note that according to the definition of $\text{RD}$, we have $\text{RD}(\tilde{x}) \in [\tilde{x}-1, \tilde{x}+1)$. Therefore, when $\text{RD}(\tilde{x}) - 2 \geq 0$, $\text{RD}(\tilde{x}) - 1 \in [1, \tilde{x})$. Then we have $E[\Delta_B(\text{RD}(\tilde{x}) - 1 - 1)] \leq E[\Delta_B(\text{RD}(\tilde{x}) - 1 + 1)]$ which implies that $E[\Delta_B(\text{RD}(\tilde{x}) - 2)] \leq E[\Delta_B(\text{RD}(\tilde{x}))]$. Similarly, when $\text{RD}(\tilde{x}) + 2 \leq n$, $\text{RD}(\tilde{x}) + 1 \in [\tilde{x}, n - 1]$, and we have $E[\Delta_B(\text{RD}(\tilde{x}))] \geq E[\Delta_B(\text{RD}(\tilde{x}) + 2)]$.

Second, we find that the properties of $\tilde{x}$ also imply that for other $x$, $E[\Delta_B(x)]$ is less than or equal to either $E[\Delta_B(\text{RD}(\tilde{x}) - 2)]$ or $E[\Delta_B(\text{RD}(\tilde{x}) + 2)]$.

Finally, when $\tilde{x} > n - 1$, for all $x < \tilde{x}$, $E[\Delta_B(x + 1)] \geq E[\Delta_B(x - 1)]$, the surprise is increasing with $x$ and $\text{RD}(\tilde{x}) = n$ is the optimal bonus.

Therefore, $E[\Delta_B(\text{RD}(\tilde{x}))]$ is the maximum and $\text{RD}(\tilde{x})$ is the optimal bonus. □

Observation 4.4  In the symmetric case, i.e. $\alpha = \beta$, $Q^i_j$ is also symmetric, that is $Q^j_i = Q^i_{1-j}$.

Proof of Observation 4.4 Recall Lemma 4.2, we calculate the value of $Q^i_j$

$$Q^i_j = \frac{2(\alpha + \beta - 1)(\binom{i}{j}(\binom{\alpha+\beta}{\alpha-1}))}{(\alpha + \beta + i)(\binom{i+\alpha+\beta}{j+\alpha})} = \frac{2(2\alpha - 1)(\binom{i}{j}(\binom{2\alpha-2}{\alpha-1}))}{(2\alpha + i)(\binom{i+2\alpha}{j+\alpha})}$$

Then we calculate the value of $Q^i_{1-j}$

$$Q^i_{1-j} = \frac{2(\alpha + \beta - 1)(\binom{i}{i-j}(\binom{\alpha+\beta}{\alpha-1}))}{(\alpha + \beta + i)(\binom{i+\alpha+\beta}{i-j+\alpha})} = \frac{2(2\alpha - 1)(\binom{i}{i-j}(\binom{2\alpha-2}{\alpha-1}))}{(2\alpha + i)(\binom{i+2\alpha}{i-j+\alpha})} = \frac{2(2\alpha - 1)(\binom{i}{j}(\binom{2\alpha-2}{\alpha-1}))}{(2\alpha + i)(\binom{i+2\alpha}{j+\alpha})}$$

Thus $Q^i_j = Q^i_{1-j}$ □