



THE PROBABILITY DISTRIBUTION OF THE NUMBER OF DISTINCT OUTCOMES IN REPEATED WEIGHTED EXPERIMENTS

CALEB BRINKMAN , LIANYONG XUE , AND TÜRKAY YOLCU *

Abstract. The number of distinct outcomes arising in repeated weighted experiments is a fundamental quantity in probability theory, with natural applications in statistics and game theory. In this paper, we pursue two complementary objectives. First, we establish explicit and recursive formulas for the probability distribution governing this quantity, then derive its expected value and variance via two independent methods: a classical combinatorial approach and a probabilistic method based on indicator random variables, joint probabilities, and covariance matrices. We demonstrate the practical relevance of this distribution through applications in weighted bootstrap sampling with minority classes and in an analysis of the NBA Draft Lottery. Second, as a central theoretical contribution, we resolve a conjecture proposed in [5] asserting that the expected value of this distribution is maximized when all outcomes are equiprobable, a result consistent with the maximum-entropy principle from information theory. We first support the conjecture through Monte Carlo simulations, which motivate and guide a complete formal proof confirming its validity.

Keywords: Probability Distribution; Weighted Experiments; Distinct Outcomes; Expected Value; Entropy.

MSC (2020): 60C05, 62E17, 91A80, 65C05.

1. INTRODUCTION

Historically, weighted probabilistic models have contributed heavily to engineering and the sciences, arising because their unweighted counterparts often lack the functionality to accurately simulate many real-world scenarios. These weighted models necessitate a concept in information theory [11] called the entropy of a random variable, which quantifies the average amount of “information,” “surprise,” or “uncertainty” associated with a random outcome. Entropy has emerged as a burgeoning concept in probability and statistics, with work by [6, 11] highlighting its capabilities. Weighted experiments also arise naturally in machine learning, where the probability of distinct outcomes is influenced by reward structures, class imbalance, or feature relevance. Applications include outcome-weighted learning for optimal decision rules [9], probability estimation for multicategory outcomes [12], and probability-weighted feature selection in high-dimensional models [1].

Received: 2025.12.21

Revised: 2026.05.02

Accepted: 2026.06.19

* Corresponding author

Caleb Brinkman (cbrin5@illinois.edu) $\diamond$ ORCID:0009-0008-7925-3647

Lianying Xue (lxue@bradley.edu) $\diamond$ ORCID:0000-0001-5896-9234

Türkay Yolcu (tyolcu@bradley.edu) $\diamond$ ORCID:0009-0002-5435-9417

In this manuscript, we explore weighted probabilistic models and entropy through the problem of determining a probability distribution for the number of distinct outcomes that occur in repeated weighted experiments. This distribution allows us to analyze how randomness behaves in weighted settings and illustrates that the highest level of entropy occurs when all outcomes are equiprobable.

To motivate our study, consider a scenario from incentive design, a construct from the Nobel Prize-winning field of contract theory. Suppose a firm manager oversees twelve employees and has twelve equal-value incentive checks to distribute at the year-end banquet. To determine the recipient of each check, the manager conducts a weighted random selection with replacement, aiming both to reward superior employees and to avoid accusations of bias. Over the year, employees who meet a universal standard earn personalized achievement cards that are placed into a jar, and during the banquet the CEO makes twelve selections with replacement from this jar to determine the recipients. Because different employees contribute different numbers of achievement cards, the selection process is conducted in a weighted and randomized manner. It is therefore pertinent to determine the probability of k distinct employees receiving at least one incentive check, for $1 \leq k \leq 12$, as well as the expected number of distinct employees that receive a check. Our model answers this question directly.

Interestingly, our results can also be employed to analyze the annual NBA Draft Lottery. The lottery determines the selection order for that year’s NBA Draft through a weighted random draw among each of the league’s non-postseason teams, where the odds of winning correlate negatively with league standing [17]. Here, we restrict our focus to the winner of the Draft Lottery, as the value of the top pick has led some teams to be accused of “tanking”, purposely losing games to increase their probability of securing the first pick [4]. In 2019, the NBA significantly overhauled their lottery odds to mitigate this issue [4], as expressed in Table 1.1. These new odds theoretically increase the diversity of teams that can obtain the top pick.

Standing	1	2	3	4	5	6	7	8	9	10	11	12	13	14
2018	0.25	0.199	0.138	0.137	0.088	0.053	0.053	0.028	0.017	0.011	0.008	0.007	0.006	0.005
2019	0.14	0.14	0.14	0.125	0.105	0.09	0.06	0.06	0.06	0.03	0.02	0.01	0.01	0.01

TABLE 1.1. Probabilities of getting the top pick in 2018 versus 2019 based on league record [18]. (1 denotes the team with the worst league standing.)

As an application of our distribution, we examine the impact of the NBA’s revised lottery odds on the randomness of the Draft by modeling the expected number of distinct pre-lottery positions that secure the top pick over a 14-year horizon (corresponding to the 14 teams in the lottery). To quantify this effect, we compute the percent increase in the expected number of distinct outcomes between the 2018 and 2019 lottery systems, which we find to be approximately 19%.

Mathematically, the NBA Draft Lottery problem represents the expected number of distinct outcomes that occur in repeated weighted experiments. We can model this event by obtaining a concise, explicit formula for a probability distribution representing this scenario, which can then be used to compute its expected value and answer our Draft Lottery question. It is fascinating how mathematics weaves its structure into everyday scenarios.

Before we embark on this journey of discovery, it is important to note that the historical foundations of our work trace back to early analyses of games of chance, beginning with Cardano's *Liber de Ludo Aleae* [2], composed around 1560 and published posthumously in 1663. Although Cardano's treatment lacked the formal structure of later developments, it introduced many of the fundamental ideas of probability in the context of gambling. These ideas were later refined by Pascal and Fermat [14] in their correspondence on equiprobable events. Huygens, in his 1657 manuscript *De Ratiociniis in Ludo Aleae* [7], synthesized and extended this earlier work, providing the first systematic exposition of expectations and unequal-probability outcomes.

In [5], the authors obtained explicit formulas for the probability distribution, expected value, and variance of the number of distinct outcomes in the setting of repeated equiprobable events (see Corollary 3.1 and Corollary 3.2). Based on these results, they further proposed a conjecture asserting that, among all possible probability assignments, the expected number of distinct outcomes is maximized in the equiprobable case.

Our work advances this line of inquiry by deriving an explicit distribution for the number of distinct outcomes produced by repeated weighted experiments, thereby generalizing the results of [5] and extending Huygens' perspective into contemporary probabilistic modeling. This explicit distribution also enables us to prove the conjecture proposed in [5].

We develop the technical setup with the following definition:

Definition 1.1. Assume that n independent trials are performed in which a random integer from 1 to n is chosen. Let the probabilities of selecting each number be $p_1, p_2, \dots, p_n$ where $\sum_{i=1}^n p_i = 1$. We define the random variable X to be the number of distinct integers that will be seen during this procedure.

As convention, we assume $\binom{0}{0} = 1$. For $k_1 + k_2 + \dots + k_m = n$, we set

$$\binom{n}{k_1, k_2, \dots, k_m} = \frac{n!}{k_1! k_2! \dots k_m!}.$$

By the Multinomial Theorem, we have

$$(p_1 + p_2 + \dots + p_m)^n = \sum_{\substack{k_1, k_2, \dots, k_m \geq 0 \\ k_1 + k_2 + \dots + k_m = n}} \binom{n}{k_1, k_2, \dots, k_m} p_1^{k_1} p_2^{k_2} \dots p_m^{k_m}.$$

2. HEURISTIC DERIVATIONS FOR THE PROBABILITY DISTRIBUTION OF X

In this section, we develop two separate derivations. Both will serve as foundational tools for establishing further theorems in the subsequent sections.

2.1. Derivations via the Complementary Principle. To derive a probability distribution for X , we leverage the *complementary principle*. In this context, we take X to be the number of distinct pre-lottery positions that receive the top draft pick, in order to provide a more tangible example for the reader's motivation.

As an example, we model a hypothetical draft lottery containing five teams performed annually over five years. We specify probabilities p_1, p_2, p_3, p_4, p_5 , where $\sum_{i=1}^5 p_i = 1$, and each p_i denotes the probability that the i -th pre-lottery position wins the lottery and receives the top draft pick.

Case: $P[X = 1]$. The probability that exactly one distinct pre-lottery position receives the top pick across all five years is

$$P[X = 1] = \sum_{1 \leq i \leq 5} p_i^5. \quad (2.1)$$

Case $P[X = 2]$. To derive $P[X = 2]$, we again use the complementary principle. Fix a pair of distinct teams (i_1, i_2) . By the Binomial Theorem, the probability that the top pick falls to two pre-lottery positions i_1 and i_2 is $(p_{i_1} + p_{i_2})^5$. Within this event, we must exclude the cases where *only one* team appears among the five picks. Using our earlier expression for $P[X = 1]$, the total probability that all five picks come from a single team in $\{i_1, i_2\}$ can be written as $p_{i_1} P[X = 1] + p_{i_2} P[X = 1]$ or $(p_{i_1} + p_{i_2}) P[X = 1]$. Thus, for this fixed pair, the probability that *exactly two* distinct teams (namely i_1 and i_2) appear is $(p_{i_1} + p_{i_2})^5 - (p_{i_1} + p_{i_2}) P[X = 1]$. Summing over all possible pairs of teams yields

$$P[X = 2] = \sum_{1 \leq i_1 < i_2 \leq 5} \left[(p_{i_1} + p_{i_2})^5 - (p_{i_1} + p_{i_2}) P[X = 1] \right].$$

The correction term scales $P[X = 1]$ by the proportion of the total probability mass concentrated in $\{i_1, i_2\}$, since $P[X = 1]$ was computed over all five positions.

Case $P[X = 3]$. To derive an expression for $P[X = 3]$, we fix a subset of three teams and isolate the probability that *exactly* these three teams appear among the five outcomes. Let $S = \{i_1, i_2, i_3\}$ be a fixed triple of distinct pre-lottery positions to which the top pick may fall.

Step 1: *At most three distinct teams from S .* The probability that all five outcomes are drawn from this triple is

$$P(\text{all draws in } S) = (p_{i_1} + p_{i_2} + p_{i_3})^5. \quad (2.2)$$

This event includes all configurations with exactly one, exactly two, or exactly three distinct teams from S .

Step 2: *Decomposition by the number of distinct teams.* Focus first on the index i_1 in S . Among the three teams in S , the number of k -element subsets that contain i_1 is $\binom{3-1}{k-1}$, since we choose the remaining $k-1$ teams from $S \setminus \{i_1\}$. Among all five teams, the number of k -element subsets that contain i_1 is $\binom{5-1}{k-1}$, corresponding to choosing the remaining $k-1$ teams from $\{1, \dots, 5\} \setminus \{i_1\}$. Thus, the fraction of k -subsets that contain i_1 and lie entirely within S is $\binom{3-1}{k-1} / \binom{5-1}{k-1}$. Consequently, the portion of the event “exactly k distinct teams occur and i_1 is among them” that is compatible with all k teams lying in S is weighted by

$$\frac{\binom{3-1}{k-1}}{\binom{5-1}{k-1}} p_{i_1} P[X = k].$$

The same reasoning applies to i_2 and i_3 .

Step 3: *Isolating exactly three distinct teams from S .*

$$P[\text{exactly 3 distinct teams from } S] = (p_{i_1} + p_{i_2} + p_{i_3})^5 - \sum_{k=1}^2 \frac{\binom{3-1}{k-1}}{\binom{5-1}{k-1}} (p_{i_1} + p_{i_2} + p_{i_3}) P[X = k].$$

Finally, summing over all triples $S = \{i_1, i_2, i_3\}$ with $1 \leq i_1 < i_2 < i_3 \leq 5$ yields

$$P[X = 3] = \sum_{1 \leq i_1 < i_2 < i_3 \leq 5} \left[(p_{i_1} + p_{i_2} + p_{i_3})^5 - \sum_{k=1}^{3-1} \frac{\binom{3-1}{k-1}}{\binom{5-1}{k-1}} (p_{i_1} + p_{i_2} + p_{i_3}) P[X = k] \right].$$

General pattern: $P[X = r]$. The structure above extends naturally.

$$P[X = r] = \sum_{1 \leq i_1 < \dots < i_r \leq 5} \left[(p_{i_1} + \dots + p_{i_r})^5 - \sum_{k=1}^{r-1} \frac{\binom{r-1}{k-1}}{\binom{5-1}{k-1}} (p_{i_1} + \dots + p_{i_r}) P[X = k] \right],$$

where the inner sum subtracts, via complementary counting, the contributions corresponding to exactly k distinct teams. The factor $\binom{r-1}{k-1} / \binom{5-1}{k-1}$ is the proportion of k -element subsets of $\{1, \dots, 5\}$ (containing a fixed index) that lie entirely within the chosen r -subset $\{i_1, \dots, i_r\}$. This weighting ensures that only the portion of $P[X = k]$ attributable to subsets contained in $\{i_1, \dots, i_r\}$ is removed, thereby isolating the probability that exactly r distinct positions are observed.

2.2. Derivations via the Inclusion-Exclusion Principle. To derive the probability distribution for X , the number of distinct pre-lottery positions that receive the top draft pick, we leverage the *Inclusion-Exclusion Principle*. As a concrete example, we model a hypothetical draft lottery containing $n = 5$ teams performed annually over 5 years. We specify probabilities p_1, p_2, p_3, p_4, p_5 , where $\sum_{i=1}^5 p_i = 1$, and each p_i denotes the probability that the i -th pre-lottery position wins the top draft pick in any given year.

To find the probability $P[X = r]$ that exactly r distinct teams win the top pick, we first choose a specific subset of r teams, denoted by $S := \{i_1, \dots, i_r\} \subset \{1, 2, 3, 4, 5\}$ where $|S| = r$. We calculate the probability that the winners over all 5 years come *only* from this subset S , and use inclusion-exclusion to ensure all r teams are chosen at least once.

Case $P[X = 1]$. The probability that exactly one distinct pre-lottery position receives the top pick across all five years is simply the sum of the probabilities of each single team winning all 5 years:

$$P[X = 1] = \sum_{1 \leq i \leq 5} p_i^5.$$

Case $P[X = 2]$. For a fixed pair of teams $\{i_1, i_2\}$, the probability that all 5 picks come from this pair is $(p_{i_1} + p_{i_2})^5$. However, this includes the cases where only i_1 wins all 5 years, or only i_2 wins all 5 years. By the Inclusion-Exclusion Principle, we subtract these single-team outcomes:

$$P[X = 2] = \sum_{1 \leq i_1 < i_2 \leq 5} [(p_{i_1} + p_{i_2})^5 - p_{i_1}^5 - p_{i_2}^5].$$

Case $P[X = 3]$. For a fixed triplet of teams $\{i_1, i_2, i_3\}$, the baseline probability that all picks fall within this triplet is $(p_{i_1} + p_{i_2} + p_{i_3})^5$. To isolate the event where *all three* appear at least once, we subtract the cases where at most two appear, and add back the cases where only

one appears:

$$P[X = 3] = \sum_{1 \leq i_1 < i_2 < i_3 \leq 5} \left[(p_{i_1} + p_{i_2} + p_{i_3})^5 - \sum_{1 \leq j < k \leq 3} (p_{i_j} + p_{i_k})^5 + \sum_{1 \leq j \leq 3} p_{i_j}^5 \right].$$

General Pattern. The structure generalizes cleanly. For any $1 \leq r \leq 5$, the probability that exactly r distinct positions are observed is found by summing over all possible subsets S of size r , and applying the complementary principle internally to each subset:

$$P[X = r] = \sum_{1 \leq i_1 < \dots < i_r \leq 5} \sum_{j=0}^{r-1} (-1)^j \sum_{\substack{T \subset \{i_1, \dots, i_r\} \\ |T|=r-j}} \left(\sum_{i \in T} p_i \right)^5.$$

Alternatively, setting $k = r - j$, we get

$$P[X = r] = \sum_{1 \leq i_1 < \dots < i_r \leq 5} \sum_{k=1}^r (-1)^{r-k} \sum_{\substack{T \subset \{i_1, \dots, i_r\} \\ |T|=k}} \left(\sum_{i \in T} p_i \right)^5. \quad (2.3)$$

3. ESTABLISHING THE PROBABILITY DISTRIBUTION OF X

Guided by the previous observations in Section 2.1, we examine the emerging patterns and use them to formulate a recursive characterization of the probability distribution of X .

Theorem 3.1. *Let N be a random variable with probability distribution*

$$P[N = k] = p_k, \quad 1 \leq k \leq n,$$

and X be the number of distinct integers observed in n independent observations of N . The probability of obtaining r distinct integers in n trials is given by

$$P[X = r] = \sum_{1 \leq i_1 < \dots < i_r \leq n} \left((p_{i_1} + \dots + p_{i_r})^n - \sum_{k=1}^{r-1} \frac{\binom{r-1}{k-1}}{\binom{n-1}{k-1}} (p_{i_1} + \dots + p_{i_r}) P[X = k] \right). \quad (3.4)$$

Proof. As shown in Section 2.1, the complementary principle yields the structure required to isolate exact cardinalities, and Theorem 3.1 follows as a direct application of this framework. To this end, let $1 \leq r \leq n$ and fix an r -subset $S := \{i_1, \dots, i_r\} \subset \{1, \dots, n\}$, and consider the event that all observed integers lie in S and that exactly r distinct integers from S appear among the n draws.

Step 1: *At most r distinct integers from S .* By the Binomial Theorem, the probability that all n draws come from S (i.e., no integer outside S appears) is

$$P[\text{all draws in } S] = (p_{i_1} + \dots + p_{i_r})^n. \quad (3.5)$$

This event includes all configurations with exactly k distinct integers from S for $1 \leq k \leq r$.

Step 2: *Decomposing by the number of distinct integers.* Let $1 \leq k \leq r$ and let $X_k(S)$ denote the event that *exactly* k distinct integers appear among the outcomes, and that all of them belong to the subset S . Then the event in (3.5) can be decomposed as

$$\{\text{all draws in } S\} = \bigcup_{k=1}^r X_k(S),$$

and these events are disjoint. Hence

$$\sum_{k=1}^r P[X_k(S)] = (p_{i_1} + \cdots + p_{i_r})^n. \quad (3.6)$$

Our goal in the decomposition above is to single out the probability $P[X_r(S)]$, which captures the event that precisely r distinct integers from S are observed.

Step 3: *Relating $P[X_k(S)]$ to $P[X = k]$.* Fix k with $1 \leq k \leq r-1$. We now express $P[X_k(S)]$ as a fraction of the global probability $P[X = k]$. Focus first on a single index $i_1 \in S$. Among the r integers in S , the number of k -element subsets that contain i_1 is $\binom{r-1}{k-1}$, since we choose the remaining $k-1$ integers from $S \setminus \{i_1\}$. Among all n integers, the number of k -element subsets that contain i_1 is $\binom{n-1}{k-1}$, corresponding to choosing the remaining $k-1$ integers from $\{1, \dots, n\} \setminus \{i_1\}$. Thus, the fraction of k -subsets (containing i_1) that lie entirely in S is $\binom{r-1}{k-1} / \binom{n-1}{k-1}$. Consequently, the portion of the event “exactly k distinct integers occur and i_1 is among them” that is compatible with all k integers lying in S is weighted by

$$\frac{\binom{r-1}{k-1}}{\binom{n-1}{k-1}} p_{i_1} P[X = k].$$

The same reasoning applies to each $i_j \in S$, $j = 1, \dots, r$. Summing over all indices in S yields

$$P[X_k(S)] = \frac{\binom{r-1}{k-1}}{\binom{n-1}{k-1}} (p_{i_1} + \cdots + p_{i_r}) P[X = k], \quad 1 \leq k \leq r-1. \quad (3.7)$$

Step 4: *Isolating exactly r distinct integers from S .* Substituting (3.7) into (3.6) gives

$$(p_{i_1} + \cdots + p_{i_r})^n = \sum_{k=1}^{r-1} \frac{\binom{r-1}{k-1}}{\binom{n-1}{k-1}} (p_{i_1} + \cdots + p_{i_r}) P[X = k] + P[X_r(S)].$$

Rearranging, we obtain

$$P[X_r(S)] = (p_{i_1} + \cdots + p_{i_r})^n - \sum_{k=1}^{r-1} \frac{\binom{r-1}{k-1}}{\binom{n-1}{k-1}} (p_{i_1} + \cdots + p_{i_r}) P[X = k].$$

That is,

$$P[\text{exactly } r \text{ distinct integers from } S] = (p_{i_1} + \cdots + p_{i_r})^n - \sum_{k=1}^{r-1} \frac{\binom{r-1}{k-1}}{\binom{n-1}{k-1}} (p_{i_1} + \cdots + p_{i_r}) P[X = k].$$

Summing over all r -subsets $S = \{i_1, \dots, i_r\}$ with $1 \leq i_1 < \cdots < i_r \leq n$ yields the desired result for $P[X = r]$. $\square$

Consider the summation of all possible $p_{i_1} + \cdots + p_{i_r}$, where $1 \leq i_1 < \cdots < i_r \leq n$. The total number of terms here is $r \binom{n}{r}$. Since each p_i appears equally, we can rearrange them to form $\frac{1}{n} [r \binom{n}{r}] = \binom{n-1}{r-1}$ blocks of $p_1 + \cdots + p_n$. Mathematically, we can express this formula as follows:

Lemma 3.1. *For any $1 \leq r \leq n$, we have*

$$\sum_{1 \leq i_1 < \cdots < i_r \leq n} (p_{i_1} + \cdots + p_{i_r}) = \binom{n-1}{r-1} (p_1 + \cdots + p_n) = \binom{n-1}{r-1}.$$

Proof. We establish the proof by mathematical induction on $n \geq 1$. The base case where $n = 1$ is simple as $n = 1$ implies that $r = 1$. Hence,

$$\sum_{1 \leq i_1 < \dots < i_r \leq 1} (p_{i_1} + \dots + p_{i_r}) = \sum_{r=1}^1 p_r = 1 = \binom{1-1}{1-1}.$$

We carry on with our inductive assumption that for some n and any $1 \leq r \leq n$,

$$\sum_{1 \leq i_1 < \dots < i_r \leq n} (p_{i_1} + \dots + p_{i_r}) = \binom{n-1}{r-1}.$$

For our inductive step with $p_1 + p_2 + \dots + p_n + p_{n+1} = 1$, let us set $q_i = \frac{p_i}{1-p_{n+1}}$ and apply the inductive hypothesis with $q_1 + q_2 + \dots + q_n = 1$ to ascertain

$$\sum_{1 \leq i_1 < \dots < i_r \leq n} (q_{i_1} + \dots + q_{i_r}) = \binom{n-1}{r-1}, \quad \sum_{1 \leq i_1 < \dots < i_{r-1} \leq n} (q_{i_1} + \dots + q_{i_{r-1}}) = \binom{n-1}{r-2}.$$

Now, we consider r -combinations of $n+1$ integers. There are two cases regarding the inclusion or exclusion of p_{n+1} .

Case 1. Let p_{n+1} be included in the sum's r -combination. We have

$$\begin{aligned} \Sigma_1 &:= \sum_{\substack{1 \leq i_1 < \dots < i_{r-1} \leq n \\ \text{including } p_{n+1}}} (p_{i_1} + \dots + p_{i_{r-1}} + p_{n+1}) \\ &= \sum_{1 \leq i_1 < \dots < i_{r-1} \leq n} (p_{i_1} + \dots + p_{i_{r-1}}) + \sum_{1 \leq i_1 < \dots < i_{r-1} \leq n} p_{n+1} \\ &= (1 - p_{n+1}) \sum_{1 \leq i_1 < \dots < i_{r-1} \leq n} (q_{i_1} + \dots + q_{i_{r-1}}) + \sum_{1 \leq i_1 < \dots < i_{r-1} \leq n} p_{n+1} \\ &= (1 - p_{n+1}) \binom{n-1}{r-2} + \binom{n}{r-1} p_{n+1}. \end{aligned}$$

Case 2. Let p_{n+1} be excluded in the sum's r -combination. We have

$$\begin{aligned} \Sigma_2 &:= \sum_{\substack{1 \leq i_1 < \dots < i_r \leq n \\ \text{excluding } p_{n+1}}} (p_{i_1} + \dots + p_{i_{r-1}} + p_{i_r}) \\ &= (1 - p_{n+1}) \sum_{1 \leq i_1 < \dots < i_r \leq n} (q_{i_1} + \dots + q_{i_{r-1}} + q_{i_r}) \\ &= (1 - p_{n+1}) \binom{n-1}{r-1}. \end{aligned}$$

Let

$$\Sigma := \sum_{1 \leq i_1 < \dots < i_r \leq n+1} (p_{i_1} + \dots + p_{i_r}) = \Sigma_1 + \Sigma_2.$$

Employing Pascal's identity,

$$\binom{n-1}{r-2} + \binom{n-1}{r-1} = \binom{n}{r-1},$$

and algebraic manipulations produces the desired result:

$$\begin{aligned}
 \Sigma &= (1 - p_{n+1}) \binom{n-1}{r-2} + \binom{n}{r-1} p_{n+1} + (1 - p_{n+1}) \binom{n-1}{r-1} \\
 &= (1 - p_{n+1}) \left[\binom{n-1}{r-2} + \binom{n-1}{r-1} \right] + \binom{n}{r-1} p_{n+1} \\
 &= (1 - p_{n+1}) \binom{n}{r-1} + \binom{n}{r-1} p_{n+1} = \binom{n}{r-1} = \binom{(n+1)-1}{r-1}.
 \end{aligned}$$

This completes the proof. $\square$

In our manuscript, Lemma 3.1 serves two purposes: simplifying our recursive formula and connecting our recursive formula to an explicit formula shown in a later theorem. Indeed, applying Lemma 3.1 after distributing the outer summation in Theorem 3.1 produces the following expression:

$$P[X = r] = \sum_{1 \leq i_1 < \dots < i_r \leq n} (p_{i_1} + \dots + p_{i_r})^n - \sum_{k=1}^{r-1} \frac{\binom{r-1}{k-1}}{\binom{n-1}{k-1}} \binom{n-1}{r-1} P[X = k]. \quad (3.8)$$

The clarity of this equation is increased by applying the popular subcommittee identity, $\binom{n}{r} \binom{r}{k} = \binom{n}{k} \binom{n-k}{r-k}$ along with the identity $\binom{m}{\ell} = \frac{m}{\ell} \binom{m-1}{\ell-1}$, which yields

$$\frac{\binom{r-1}{k-1}}{\binom{n-1}{k-1}} \binom{n-1}{r-1} = \frac{\binom{n-1}{r-1} \binom{r-1}{k-1}}{\binom{n-1}{k-1}} = \frac{\binom{n-1}{k-1} \binom{n-k}{r-k}}{\binom{n-1}{k-1}} = \binom{n-k}{r-k},$$

which can be substituted into (3.8) to obtain:

$$P[X = r] = \sum_{1 \leq i_1 < \dots < i_r \leq n} (p_{i_1} + \dots + p_{i_r})^n - \sum_{k=1}^{r-1} \binom{n-k}{r-k} P[X = k]. \quad (3.9)$$

To demonstrate the second use of Lemma 3.1, we first prove the following fact, which results in an explicit formula for the probability distribution of X . While this should signal the end of our search, we continue on, believing this explicit formula's complexity makes calculating the expected value and variance unnecessarily difficult. Though this explicit formula lacks the brevity we desire, its value stems from its connection to (3.9), which will prove invaluable to confirming our recursive formula's truth.

Theorem 3.2. *Let N be a random variable with probability distribution*

$$P[N = k] = p_k, \quad 1 \leq k \leq n,$$

and X be the number of distinct integers observed in n independent observations of N . Then the probability distribution of X is given by

$$P[X = r] = \sum_{1 \leq i_1 < \dots < i_r \leq n} \left[\sum_{\substack{y_{i_1} + \dots + y_{i_r} = n \\ \text{all } y_{i_j} \geq 1}} \frac{n!}{y_{i_1}! \dots y_{i_r}!} p_{i_1}^{y_{i_1}} \dots p_{i_r}^{y_{i_r}} \right], \quad r = 1, 2, \dots, n. \quad (3.10)$$

Proof. Let $Y_k = B(n, p_k)$ be a binomial random variable expressing the number of times that k occurs in the n observations. Then $Y = (Y_1, \dots, Y_n)$ has a multinomial distribution

with probabilities $p_1, \dots, p_n$, number of observations n , and joint probability function

$$p(y) = p(y_1, \dots, y_n) = \binom{n}{y_1, \dots, y_n} p_1^{y_1} \cdots p_n^{y_n} = \frac{n!}{y_1! \cdots y_n!} p_1^{y_1} \cdots p_n^{y_n}$$

where $y_1 + \cdots + y_n = n$. We now define the vectors

$$e_1 = (1, 0, \dots, 0) \quad e_2 = (0, 1, \dots, 0) \quad \dots \quad e_n = (0, 0, \dots, 1), \quad (3.11)$$

These vectors enable us to alternatively represent our joint probability function as

$$p(y_1, \dots, y_n) = p(y_1 e_1 + \cdots + y_n e_n) = \frac{n!}{y_1! \cdots y_n!} p_1^{y_1} \cdots p_n^{y_n}.$$

Hence, for $r = 1, \dots, n$, we have

$$\begin{aligned} P[X = r] &= \sum_{1 \leq i_1 < \cdots < i_r \leq n} \left[\sum_{\substack{y_{i_1} + \cdots + y_{i_r} = n \\ \text{all } y_{i_j} \geq 1}} p(y_{i_1} e_{i_1} + \cdots + y_{i_r} e_{i_r}) \right] \\ &= \sum_{1 \leq i_1 < \cdots < i_r \leq n} \left[\sum_{\substack{y_{i_1} + \cdots + y_{i_r} = n \\ \text{all } y_{i_j} \geq 1}} \frac{n!}{y_{i_1}! \cdots y_{i_r}!} p_{i_1}^{y_{i_1}} \cdots p_{i_r}^{y_{i_r}} \right], \end{aligned}$$

which is exactly what we desired to prove. $\square$

The connection between (3.9) and Theorem 3.2 can be observed in Remark 3.1 below.

Remark 3.1. We observe the following probabilities for $P[X = r]$:

$$\begin{aligned} (a) \quad r = 1: \quad P[X = 1] &= \sum_{k=1}^n p(n e_k) = \sum_{k=1}^n p_k^n. \\ (b) \quad r = 2: \quad P[X = 2] &= \sum_{1 \leq i < j \leq n} \left[\sum_{\substack{y_i + y_j = n \\ y_i \geq 1, y_j \geq 1}} p(y_i e_i + y_j e_j) \right] \\ &= \sum_{1 \leq i < j \leq n} \left[\left(\sum_{y_i + y_j = n} p(y_i e_i + y_j e_j) \right) - p(n e_i) - p(n e_j) \right] \\ &= \sum_{1 \leq i < j \leq n} \left[\left(\sum_{y_i + y_j = n} \frac{n!}{y_i! y_j!} p_i^{y_i} p_j^{y_j} \right) - p_i^n - p_j^n \right] \\ &= \sum_{1 \leq i < j \leq n} (p_i + p_j)^n - \sum_{1 \leq i < j \leq n} (p_i^n + p_j^n) \\ &= \sum_{1 \leq i < j \leq n} (p_i + p_j)^n - (n-1)P[X = 1]. \\ (c) \quad r = n: \quad P[X = n] &= p(1, \dots, 1) = n!(p_1 \cdots p_n). \end{aligned}$$

We highlight the connection between our recursive and explicit formulas through the calculation of $P[X = 2]$. Note how the vectorial representation in (3.11) clarifies the terms taken out of the inner summation. From this, we apply Lemma 3.1 to acquire our final result, connecting the findings from Theorem 3.2 to our recursive formula in (3.9). We leverage this connection to verify our recursive formula with the ensuing theorem.

Theorem 3.3. *Let N be a random variable with probability distribution*

$$P[N = k] = p_k, \quad 1 \leq k \leq n,$$

and X be the number of distinct integers observed in n independent observations of N . We have the following recursive formula for the probabilities:

$$P[X = r] = \sum_{1 \leq i_1 < \dots < i_r \leq n} (p_{i_1} + \dots + p_{i_r})^n - \sum_{k=1}^{r-1} \binom{n-k}{r-k} P[X = k]. \quad (3.12)$$

Proof. We establish this result in two parts: the explicit formula provided by Theorem 3.2 and the complementary counting method used to derive Theorem 3.1. To find $P[X = r]$, we leverage the multinomial theorem to calculate $P[X = s]$ for any s in $\{1, 2, \dots, r\}$. We then use the explicit formula from Theorem 3.2 to subtract the probability $P[X = t]$ for any t in $\{1, 2, \dots, r-1\}$. Calculating this for every r -combination in the set $\{1, 2, \dots, n\}$ gives the expression

$$P[X = r] = \sum_{1 \leq i_1 < \dots < i_r \leq n} \sum_{\substack{y_{i_1} + \dots + y_{i_r} = n \\ \text{all } y_{i_j} \geq 1}} \frac{n!}{y_{i_1}! \dots y_{i_r}!} p_{i_1}^{y_{i_1}} \dots p_{i_r}^{y_{i_r}} := P_1 - P_2$$

where

$$P_1 = \sum_{1 \leq i_1 < \dots < i_r \leq n} \sum_{y_{i_1} + \dots + y_{i_r} = n} \frac{n!}{y_{i_1}! \dots y_{i_r}!} p_{i_1}^{y_{i_1}} \dots p_{i_r}^{y_{i_r}} = \sum_{1 \leq i_1 < \dots < i_r \leq n} (p_{i_1} + \dots + p_{i_r})^n,$$

and

$$P_2 = \sum_{1 \leq i_1 < \dots < i_r \leq n} \sum_{k=1}^{r-1} \sum_{1 \leq j_1 < \dots < j_{r-k} \leq r} \sum_{\substack{y_{i_1} + \dots + y_{i_r} = n \\ y_{j_1} = \dots = y_{j_{r-k}} = 0}} \frac{n!}{y_{i_1}! \dots y_{i_r}!} p_{i_1}^{y_{i_1}} \dots p_{i_r}^{y_{i_r}}.$$

While one may be fairly familiar with P_1 from our recursive formula, P_2 may seem daunting at first glance due to its four summations. To simplify P_2 , we first swap the outermost two sums to obtain

$$P_2 = \sum_{k=1}^{r-1} \sum_{1 \leq i_1 < \dots < i_r \leq n} \sum_{1 \leq j_1 < \dots < j_{r-k} \leq r} \sum_{\substack{y_{i_1} + \dots + y_{i_r} = n \\ y_{j_1} = \dots = y_{j_{r-k}} = 0}} \frac{n!}{y_{i_1}! \dots y_{i_r}!} p_{i_1}^{y_{i_1}} \dots p_{i_r}^{y_{i_r}}.$$

We now use (3.10) to turn two sums into one, yielding

$$P_2 = \sum_{k=1}^{r-1} \binom{n}{r} \binom{r}{r-k} \left[\frac{1}{\binom{n}{k}} \sum_{1 \leq l_1 < \dots < l_k \leq n} \sum_{\substack{y_{l_1} + \dots + y_{l_k} = n \\ \text{all } y_{l_j} \geq 1}} \frac{n!}{y_{l_1}! \dots y_{l_k}!} p_{l_1}^{y_{l_1}} \dots p_{l_k}^{y_{l_k}} \right].$$

To obtain $P[X = k]$ in the above without over-counting, it is necessary to divide by $\binom{n}{k}$. Next, we apply (3.10) and the subcommittee identity to P_2 , which produces the expression

$$P_2 = \sum_{k=1}^{r-1} \frac{\binom{n}{r} \binom{r-k}{r-k}}{\binom{n}{k}} P[X = k] = \sum_{k=1}^{r-1} \frac{\binom{n}{k} \binom{n-k}{r-k}}{\binom{n}{k}} P[X = k] = \sum_{k=1}^{r-1} \binom{n-k}{r-k} P[X = k].$$

Putting P_1 and P_2 together, we get our desired result:

$$P[X = r] = \sum_{1 \leq i_1 < \dots < i_r \leq n} (p_{i_1} + \dots + p_{i_r})^n - \sum_{k=1}^{r-1} \binom{n-k}{r-k} P[X = k].$$

This completes the proof of Theorem 3.3. $\square$

Remark 3.2. Lemma 3.1 furnishes the link between Theorem 3.1 and Theorem 3.3, thereby yielding an alternative proof of Theorem 3.3 via Theorem 3.1.

Because the laborious nature of the recursive formula makes it unsuitable for actual calculations, we instead utilize it to derive a clear, explicit formula that enables us to calculate $P[X = r]$ directly for $1 \leq r \leq n$. For brevity, we let

$$Q(n, k) := \sum_{1 \leq i_1 < \dots < i_k \leq n} (p_{i_1} + \dots + p_{i_k})^n. \quad (3.13)$$

We now embark on our journey of discovery by employing (3.9) to calculate $P[X = r]$ for $n = 5$ and $1 \leq r \leq 5$, leading to the pattern:

$$\begin{aligned} P[X = 1] &= \binom{4}{0} Q(5, 1), \\ P[X = 2] &= -\binom{4}{1} Q(5, 1) + \binom{3}{0} Q(5, 2), \\ P[X = 3] &= \binom{4}{2} Q(5, 1) - \binom{3}{1} Q(5, 2) + \binom{2}{0} Q(5, 3), \\ P[X = 4] &= -\binom{4}{3} Q(5, 1) + \binom{3}{2} Q(5, 2) - \binom{2}{1} Q(5, 3) + \binom{1}{0} Q(5, 4), \\ P[X = 5] &= \binom{4}{4} Q(5, 1) - \binom{3}{3} Q(5, 2) + \binom{2}{2} Q(5, 3) - \binom{1}{1} Q(5, 4) + \binom{0}{0} Q(5, 5). \end{aligned}$$

A careful observation of the above findings culminates our search with a simple explicit formula for the probability distribution of X .

Theorem 3.4. Let N be a random variable with probability distribution

$$P[N = k] = p_k, \quad 1 \leq k \leq n,$$

and X be the number of distinct integers observed in n independent observations of N . The probability of obtaining r distinct integers in n trials is explicitly given by

$$P[X = r] = \sum_{k=1}^r (-1)^{r-k} \binom{n-k}{r-k} \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} (p_{i_1} + \dots + p_{i_k})^n \right). \quad (3.14)$$

Proof of Theorem 3.4 by Induction. To prove this result, we shall use induction on r and Theorem 3.3. As in (3.13), we use

$$Q(n, k) = \sum_{1 \leq i_1 < \dots < i_k \leq n} (p_{i_1} + \dots + p_{i_k})^n.$$

Note that $\binom{n-1}{0} = 1$. When $r = 1$, (3.12) becomes

$$P[X = 1] = Q(n, 1) = \sum_{k=1}^1 (-1)^{1-k} \binom{n-k}{1-k} Q(n, k),$$

which is consistent with (3.14). Now assume that

$$P[X = m] = \sum_{k=1}^m (-1)^{m-k} \binom{n-k}{m-k} Q(n, k) \quad (3.15)$$

holds true for any m with $1 \leq m \leq r < n$. Let's prove this for $m = r + 1 \leq n$. Recall

$$\sum_{t=1}^r \left[\sum_{k=1}^t a(k, t) \right] = \sum_{k=1}^r \left[\sum_{t=k}^r a(k, t) \right] = \sum_{k=1}^r \left[\sum_{t=0}^{r-k} a(k, t+k) \right]. \quad (3.16)$$

Another well-known fact from Pascal's Triangle is the symmetry property:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{n}{n-k}. \quad (3.17)$$

Furthermore, using the subcommittee identity, we have

$$\binom{n-k}{n-i} \binom{n-i}{n-r-1} = \binom{n-k}{n-r-1} \binom{r+1-k}{r+1-i} \quad (3.18)$$

for integers $n \geq r + 1 \geq i \geq k \geq 0$. Finally, using the Binomial Theorem we get

$$\sum_{i=0}^{r+1-k} (-1)^i \binom{r+1-k}{i} = (1-1)^{r+1-k} = 0,$$

and hence we arrive at

$$\sum_{i=0}^{r-k} (-1)^i \binom{r+1-k}{i} = -(-1)^{r+1-k} \binom{r+1-k}{r+1-k}. \quad (3.19)$$

Also, we can write

$$Q(n, r+1) = (-1)^{r+1-(r+1)} \binom{n-(r+1)}{r+1-(r+1)} Q(n, r+1). \quad (3.20)$$

Using Theorem 3.3 and taking into account the inductive hypothesis in (3.15) together with the identities in (3.16), (3.17), (3.18), (3.19) and (3.20), we observe that

$$\begin{aligned} P[X = r+1] &= \sum_{1 \leq i_1 < \dots < i_{r+1} \leq n} (p_{i_1} + \dots + p_{i_{r+1}})^n - \sum_{i=1}^{(r+1)-1} \binom{n-i}{(r+1)-i} P[X = i] \\ &\stackrel{(3.17)}{=} Q(n, r+1) - \sum_{i=1}^r \binom{n-i}{n-r-1} \left[\sum_{k=1}^i (-1)^{i-k} \binom{n-k}{n-i} Q(n, k) \right] \\ &\stackrel{(3.16) \& (3.18)}{=} Q(n, r+1) - \sum_{k=1}^r Q(n, k) \binom{n-k}{n-r-1} \left[\sum_{i=k}^r (-1)^{i-k} \binom{r+1-k}{r+1-i} \right] \\ &\stackrel{(3.16) \& (3.17)}{=} Q(n, r+1) - \sum_{k=1}^r Q(n, k) \binom{n-k}{n-r-1} \left[\sum_{i=0}^{r-k} (-1)^i \binom{r+1-k}{i} \right] \\ &\stackrel{(3.19) \& (3.20)}{=} \sum_{k=1}^{r+1} (-1)^{r+1-k} \binom{n-k}{r+1-k} Q(n, k) \\ &= \sum_{k=1}^{(r+1)} (-1)^{(r+1)-k} \binom{n-k}{(r+1)-k} \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} (p_{i_1} + \dots + p_{i_k})^n \right), \end{aligned}$$

which completes the proof by induction. $\square$

We now present an alternative proof of the explicit formula in (3.14). This argument relies on the inclusion-exclusion principle introduced in Section 2.2. Since the steps closely parallel those already given, we leave the details to the reader.

Theorem 3.5. *Let N be a random variable with probability distribution*

$$P[N = k] = p_k, \quad 1 \leq k \leq n,$$

and X be the number of distinct integers observed in n independent observations of N . The probability of obtaining r distinct integers in n trials is given by

$$P[X = r] = \sum_{1 \leq i_1 < \dots < i_r \leq n} \sum_{k=1}^r (-1)^{r-k} \sum_{\substack{T \subset \{i_1, \dots, i_r\} \\ |T|=k}} \left(\sum_{i \in T} p_i \right)^n. \quad (3.21)$$

Proof of Theorem 3.4 by the Inclusion-Exclusion Principle. Fix k with $0 \leq k \leq r \leq n$. Let

$$J = \{1, \dots, n\} \quad \text{and} \quad S = \{i_1, \dots, i_r\} \subset J.$$

Note that we can write

$$\sum_{1 \leq i_1 < \dots < i_r \leq n} \sum_{\substack{T \subset \{i_1, \dots, i_r\} \\ |T|=k}} \left(\sum_{i \in T} p_i \right)^n = \sum_{\substack{S \subset J \\ |S|=r}} \sum_{\substack{T \subset S \\ |T|=k}} \left(\sum_{i \in T} p_i \right)^n \quad (3.22)$$

The double sum $\sum_{\substack{S \subset J \\ |S|=r}} \sum_{\substack{T \subset S \\ |T|=k}}$ ranges over all pairs (S, T) with $T \subset S \subset J$, $|T| = k$, and $|S| = r$. Interchanging the order of summation, and grouping these pairs by T first gives

$$\sum_{\substack{S \subset J \\ |S|=r}} \sum_{\substack{T \subset S \\ |T|=k}} \left(\sum_{i \in T} p_i \right)^n = \sum_{\substack{T \subset J \\ |T|=k}} \left(\sum_{i \in T} p_i \right)^n \left| \{ S : T \subset S \subset J, |S| = r \} \right|. \quad (3.23)$$

Each pair (S, T) is counted exactly once on either side. For a fixed T with $|T| = k$, an admissible S is obtained by adjoining $r - k$ further elements chosen from $J \setminus T$, a set of size $n - k$. Hence

$$\left| \{ S : T \subset S \subset J, |S| = r \} \right| = \binom{n-k}{r-k}, \quad (3.24)$$

independently of T . Therefore, combining (3.22), (3.23), and (3.24), we have

$$\sum_{1 \leq i_1 < \dots < i_r \leq n} \sum_{\substack{T \subset \{i_1, \dots, i_r\} \\ |T|=k}} \left(\sum_{i \in T} p_i \right)^n = \binom{n-k}{r-k} \sum_{\substack{T \subset J \\ |T|=k}} \left(\sum_{i \in T} p_i \right)^n. \quad (3.25)$$

Using (3.21) together with (3.25), we obtain

$$P[X = r] = \sum_{k=1}^r (-1)^{r-k} \binom{n-k}{r-k} \sum_{\substack{T \subset J \\ |T|=k}} \left(\sum_{i \in T} p_i \right)^n.$$

Writing the inner sum $\sum_{\substack{T \subset J \\ |T|=k}}$ explicitly over all k -tuples $1 \leq i_1 < \dots < i_k \leq n$ yields (3.14). $\square$

Remark 3.3. Yet another way of seeing (3.25) is to regroup the terms. Let $J = \{1, \dots, n\}$ and $S = \{i_1, \dots, i_r\} \subset J$, and fix k with $0 \leq k \leq r$. We shall count the terms of the form $(\sum_{i \in T} p_i)^n$. Summing over every admissible pair (first choosing an r -subset $S \subset J$, then a k -subset $T \subset S$) produces $\binom{n}{r} \binom{r}{k}$ terms of the form $(\sum_{i \in T} p_i)^n$, counted with repetition, since the same T can arise from several different S . Rather than collecting these terms by S , that is, into the inner sums

$$\sum_{\substack{T \subset \{i_1, \dots, i_r\} \\ |T|=k}} \left(\sum_{i \in T} p_i \right)^n,$$

each containing $\binom{r}{k}$ terms, we collect them into copies of the full sum

$$\sum_{\substack{T \subset J \\ |T|=k}} \left(\sum_{i \in T} p_i \right)^n,$$

each containing $\binom{n}{k}$ terms. By symmetry every k -subset $T \subset J$ occurs equally often, so the terms assemble into whole copies of this full sum. The number of copies is therefore

$$\frac{\binom{n}{r} \binom{r}{k}}{\binom{n}{k}},$$

and consequently

$$\sum_{\substack{S \subset J \\ |S|=r}} \sum_{\substack{T \subset S \\ |T|=k}} \left(\sum_{i \in T} p_i \right)^n = \frac{\binom{n}{r} \binom{r}{k}}{\binom{n}{k}} \sum_{\substack{T \subset J \\ |T|=k}} \left(\sum_{i \in T} p_i \right)^n.$$

Using the identity $\binom{n}{r} \binom{r}{k} = \binom{n}{k} \binom{n-k}{r-k}$ together with (3.22), we arrive at (3.25).

Remark 3.4. Observe that if in particular $p_i = 1/n$ for all $i = 1, \dots, n$, then

$$\sum_{1 \leq i_1 < \dots < i_k \leq n} (p_{i_1} + \dots + p_{i_k})^n = \sum_{1 \leq i_1 < \dots < i_k \leq n} \left(\frac{k}{n} \right)^n = \binom{n}{k} \left(\frac{k}{n} \right)^n.$$

Using Remark 3.4, we can match our results with the findings in [5] as follows.

Corollary 3.1. If n events are independent and equally likely to occur (i.e., $p_i = 1/n$ for all $i = 1, \dots, n$), then

$$P[X = r] = \sum_{k=1}^r (-1)^{r-k} \binom{n-k}{r-k} \left[\binom{n}{k} \left(\frac{k}{n} \right)^n \right].$$

Theorem 3.4 simplifies calculating the expected value and variance of X , which will be utilized to answer the Draft Lottery question posed at the start of this paper.

Theorem 3.6. Let N be a random variable with probability distribution

$$P[N = k] = p_k, \quad 1 \leq k \leq n,$$

and X be the number of distinct integers observed in n observations of N with repetition. Then

(a) The expected value of X is given by

$$E(X) = \sum_{k=1}^n (1 - (1 - p_k)^n) = n - \sum_{k=1}^n (1 - p_k)^n.$$

(b) The variance of X is given by

$$\text{Var}(X) = \sum_{k=1}^n (1 - p_k)^n - (1 - p_k)^{2n} + 2 \sum_{1 \leq j < k \leq n} (1 - p_j - p_k)^n - (1 - p_j)^n (1 - p_k)^n.$$

Proof of (a). We will provide two proofs: one using the probability distribution from Theorem 3.4 along with combinatorial identities, and one using the indicator random variable, which proves more efficient in deriving the distribution's expected value.

Proof 1. Here we use Theorem 3.4 to derive the expected value formula:

$$P[X = r] = \sum_{k=1}^r (-1)^{r-k} \binom{n-k}{r-k} Q(n, k), \quad Q(n, k) := \sum_{1 \leq i_1 < \dots < i_k \leq n} (p_{i_1} + \dots + p_{i_k})^n.$$

Employing Theorem 3.4, we can rewrite $E(X) = \sum_{r=1}^n r P[X = r]$ as:

$$E(X) = \sum_{r=1}^n (n - (n - r)) \left(\sum_{k=1}^r (-1)^{r-k} \binom{n-k}{r-k} Q(n, k) \right).$$

Upon applying the distributive property and the fact that $n = \sum_{r=1}^n n P[X = r]$ yields

$$E(X) = n - \sum_{r=1}^{n-1} (n - r) \left(\sum_{k=1}^r (-1)^{r-k} \binom{n-k}{n-r} Q(n, k) \right).$$

Since $\binom{n-k}{n-r} = \frac{n-k}{n-r} \binom{n-1-k}{n-1-r}$, re-indexing the inner summation enables us to obtain

$$E(X) = n - \sum_{k=1}^{n-1} (n - k) Q(n, k) \left(\sum_{r=k}^{n-1} (-1)^{r-k} \binom{n-1-k}{n-1-r} \right).$$

The following identity then allows for the removal of the inner summation:

$$\sum_{r=k}^{n-1} (-1)^{r-k} \binom{n-1-k}{n-1-r} = \begin{cases} 0 & \text{if } k < n-1 \\ 1 & \text{if } k = n-1 \end{cases} := \delta_k(n-1),$$

which allows us to reduce the expected value formula to

$$E(X) = n - \sum_{k=1}^{n-1} (n - k) Q(n, k) \delta_k(n-1) = n - (n - (n-1)) Q(n, n-1).$$

Finally, utilizing the equality

$$Q(n, n-1) = \sum_{1 \leq i_1 < \dots < i_{n-1} \leq n} (p_{i_1} + \dots + p_{i_{n-1}})^n = \sum_{k=1}^n (1 - p_k)^n$$

provides our desired result for the expected value of X :

$$E(X) = \sum_{k=1}^n [1 - (1 - p_k)^n].$$

Proof 2. Let I_k be the indicator random variable in n observations of N . Then

$$X = \sum_{k=1}^n I_k.$$

Let $Y_k = B(n, p_k)$, the number of times that k occurs in n observations. Thus, we obtain $P(I_k = 0) = P(Y_k = 0) = (1 - p_k)^n$ and $P(I_k = 1) = 1 - (1 - p_k)^n$. Hence, we acquire $E(I_k) = 1 - (1 - p_k)^n$, allowing us to arrive at the desired result

$$E(X) = E\left(\sum_{k=1}^n I_k\right) = \sum_{k=1}^n E(I_k) = \sum_{k=1}^n [1 - (1 - p_k)^n].$$

Proof of (b). Similar to (a), we will provide two proofs: one utilizing combinatorial arguments and the other utilizing the indicator random variable.

Proof 1. Using the identity in [5],

$$\sum_{r=0}^m (-1)^r r^p \binom{m}{r} = 0 \quad \text{if } 0 \leq p \leq m-1, \quad m \geq 1,$$

we uncover the following three results that we will employ to derive $E(X^2)$.

$$\sum_{r=0}^m (-1)^r \binom{m}{r} = \begin{cases} 1 & \text{if } m = 0 \\ 0 & \text{if } m \geq 1 \end{cases}, \quad \sum_{r=0}^m (-1)^r r \binom{m}{r} = \begin{cases} 0 & \text{if } m = 0 \\ -1 & \text{if } m = 1 \\ 0 & \text{if } m \geq 2 \end{cases}, \quad (3.26)$$

$$\sum_{r=0}^m (-1)^r r^2 \binom{m}{r} = \begin{cases} 0 & \text{if } m = 0 \\ -1 & \text{if } m = 1 \\ 2 & \text{if } m = 2 \\ 0 & \text{if } m \geq 3 \end{cases}. \quad (3.27)$$

We begin by observing that

$$E(X^2) = \sum_{r=1}^n r^2 P[X = r] = \sum_{r=1}^n r^2 \left(\sum_{k=1}^r (-1)^{r-k} \binom{n-k}{r-k} Q(n, k) \right).$$

Applying the idea from (3.16) to rearrange the double summation then yields:

$$E(X^2) = \sum_{k=1}^n Q(n, k) \left(\sum_{r=0}^{n-k} (-1)^r (r+k)^2 \binom{n-k}{r} \right).$$

Whence expansion of terms allows us to get

$$\begin{aligned} E(X^2) &= \sum_{k=1}^n Q(n, k) \left(\sum_{r=0}^{n-k} (-1)^r r^2 \binom{n-k}{r} \right) + \sum_{k=1}^n 2k Q(n, k) \left(\sum_{r=0}^{n-k} (-1)^r r \binom{n-k}{r} \right) \\ &\quad + \sum_{k=1}^n k^2 Q(n, k) \left(\sum_{r=0}^{n-k} (-1)^r \binom{n-k}{r} \right). \end{aligned}$$

Use of (3.26) and (3.27) allows us to conclude that

$$\begin{aligned} E(X^2) &= Q(n, n-2) \left(\sum_{r=0}^2 (-1)^r r^2 \binom{2}{r} \right) + Q(n, n-1) \left(\sum_{r=0}^1 (-1)^r r^2 \binom{2}{r} \right) \\ &\quad + 2(n-1) Q(n, n-1) \left(\sum_{r=0}^1 (-1)^r r \binom{1}{r} \right) + n^2 Q(n, n) \left(\sum_{r=0}^0 (-1)^r \binom{0}{r} \right). \end{aligned}$$

Invoking (3.26) and (3.27) again produces

$$E(X^2) = n^2 Q(n, n) - (2n-1)Q(n, n-1) + 2Q(n, n-2).$$

Knowing that $Q(n, n) = 1$, $Q(n, n-1) = \sum_{k=1}^n (1-p_k)^n$, and

$$Q(n, n-2) = \sum_{1 \leq i_1 < \dots < i_{n-2} \leq n} (p_{i_1} + \dots + p_{i_{n-2}})^n = \sum_{1 \leq j < k \leq n} (1-p_j - p_k)^n,$$

we arrive at

$$E(X^2) = n^2 + (1-2n) \sum_{k=1}^n (1-p_k)^n + 2 \sum_{1 \leq j < k \leq n} (1-p_j - p_k)^n.$$

Additionally, we can use the expected value formula derived in (a) to say that

$$\begin{aligned} E(X)^2 &= \left(n - \sum_{k=1}^n (1-p_k)^n \right)^2 = n^2 - 2n \sum_{k=1}^n (1-p_k)^n + \left(\sum_{k=1}^n (1-p_k)^n \right)^2 \\ &= n^2 - 2n \sum_{k=1}^n (1-p_k)^n + \sum_{k=1}^n (1-p_k)^{2n} + 2 \sum_{1 \leq j < k \leq n} (1-p_j)^n (1-p_k)^n. \end{aligned}$$

With the formulas for $E(X^2)$ and $E(X)^2$ calculated, we are able to derive the variance of X as follows:

$$\begin{aligned} \text{Var}(X) &= E(X^2) - E(X)^2 \\ &= \sum_{k=1}^n [(1-p_k)^n - (1-p_k)^{2n}] + 2 \sum_{1 \leq j < k \leq n} [(1-p_j - p_k)^n - (1-p_j)^n (1-p_k)^n]. \end{aligned}$$

Proof 2. Using the indicator random variable, we observe $P(I_k = 0) = (1-p_k)^n$ and

$$P(I_j = 0 \text{ and } I_k = 0) = (1-p_j - p_k)^n \text{ for any } 1 \leq j < k \leq n.$$

Since $P(I_k = 0) = P(I_j = 0 \text{ and } I_k = 0) + P(I_j = 1 \text{ and } I_k = 0)$, we acquire

$$\begin{aligned} P(I_j = 1 \text{ and } I_k = 0) &= P(I_k = 0) - P(I_j = 0 \text{ and } I_k = 0) \\ &= (1-p_k)^n - (1-p_j - p_k)^n. \end{aligned}$$

Similarly, we can calculate the probabilities

$$\begin{aligned} P(I_j = 1, I_k = 0) &= (1-p_k)^n - (1-p_j - p_k)^n, \\ P(I_j = 1, I_k = 1) &= 1 - (1-p_j)^n - (1-p_k)^n + (1-p_j - p_k)^n, \end{aligned}$$

and utilize them to construct the joint probability table of (I_j, I_k) below.

	$I_k = 0$	$I_k = 1$	$P(I_j)$
$I_j = 0$	$(1 - p_j - p_k)^n$	$(1 - p_j)^n - (1 - p_j - p_k)^n$	$(1 - p_j)^n$
$I_j = 1$	$(1 - p_k)^n - (1 - p_j - p_k)^n$	$1 - (1 - p_j)^n - (1 - p_k)^n + (1 - p_j - p_k)^n$	$1 - (1 - p_j)^n$
$P(I_k)$	$(1 - p_k)^n$	$1 - (1 - p_k)^n$	1

Employing $E(I_k) = 1 - (1 - p_k)^n$, we procure $E(I_j I_k)$, $\text{Cov}(I_j, I_k)$, and $\text{Cov}(I_k, I_k)$:

$$E(I_j I_k) = \sum_{I_j=0}^1 \sum_{I_k=0}^1 (I_j I_k) P(I_j, I_k) = 1 - (1 - p_j)^n - (1 - p_k)^n + (1 - p_j - p_k)^n,$$

$$\text{Cov}(I_j, I_k) = E(I_j I_k) - E(I_j)E(I_k) = (1 - p_j - p_k)^n - (1 - p_j)^n(1 - p_k)^n,$$

$$\text{Cov}(I_k, I_k) = \text{Var}(I_k) = E(I_k^2) - E(I_k)^2 = (1 - p_k)^n - (1 - p_k)^{2n}.$$

Using the findings above, we acquire the variance $\text{Var}(X) = \text{Var}(\sum_{k=1}^n I_k)$ of X :

$$\begin{aligned} \text{Var}(X) &= \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \text{Cov}(I_1, I_1) & \text{Cov}(I_1, I_2) & \cdots & \text{Cov}(I_1, I_n) \\ \text{Cov}(I_2, I_1) & \text{Cov}(I_2, I_2) & \cdots & \text{Cov}(I_2, I_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(I_n, I_1) & \text{Cov}(I_n, I_2) & \cdots & \text{Cov}(I_n, I_n) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \\ &= \sum_{k=1}^n \text{Cov}(I_k, I_k) + 2 \sum_{1 \leq j < k \leq n} \text{Cov}(I_j, I_k) \\ &= \sum_{k=1}^n [(1 - p_k)^n - (1 - p_k)^{2n}] + 2 \sum_{1 \leq j < k \leq n} [(1 - p_j - p_k)^n - (1 - p_j)^n(1 - p_k)^n], \end{aligned}$$

as desired. This completes the proof. $\square$

Theorem 3.6 leads to the subsequent corollary previously obtained in [5].

Corollary 3.2. *Let N be a random variable with probability distribution*

$$P[N = i] = \frac{1}{n}, \quad 1 \leq i \leq n,$$

and X be the number of distinct integers that occur in n observations of N with repetition. Then

$$\begin{aligned} (a) \quad E(X) &= n \left[1 - \left(1 - \frac{1}{n} \right)^n \right]. \\ (b) \quad \text{Var}(X) &= n(n-1) \left(\frac{n-2}{n} \right)^n + n \left(\frac{n-1}{n} \right)^n - n^2 \left(\frac{n-1}{n} \right)^{2n}. \end{aligned}$$

Proof of (a). Setting $p_k = 1/n$ for $1 \leq k \leq n$ in part (a) of Theorem 3.6 yields

$$E(X) = \sum_{k=1}^n [(1 - (1 - p_k)^n)] = \sum_{k=1}^n \left[1 - \left(1 - \frac{1}{n} \right)^n \right] = n \left[1 - \left(1 - \frac{1}{n} \right)^n \right].$$

Proof of (b). Setting $p_k = 1/n$ for $1 \leq k \leq n$ in part (b) of Theorem 3.6 gives

$$\begin{aligned} \text{Var}(X) &= \sum_{k=1}^n [(1 - p_k)^n - (1 - p_k)^{2n}] + 2 \sum_{1 \leq j < k \leq n} [(1 - p_j - p_k)^n - (1 - p_j)^n (1 - p_k)^n] \\ &= n \left[\left(1 - \frac{1}{n}\right)^n - \left(1 - \frac{1}{n}\right)^{2n} \right] + 2 \binom{n}{2} \left[\left(1 - \frac{1}{n} - \frac{1}{n}\right)^n - \left(1 - \frac{1}{n}\right)^n \left(1 - \frac{1}{n}\right)^n \right] \\ &= n(n-1) \left(\frac{n-2}{n}\right)^n + n \left(\frac{n-1}{n}\right)^n - n^2 \left(\frac{n-1}{n}\right)^{2n}, \end{aligned}$$

as desired. $\square$

4. APPLICATIONS

4.1. Weighted Minority Representation in Bootstrap Sampling. The distribution of distinct outcomes plays a meaningful role in analyzing the behavior of bootstrap sampling procedures commonly used in ensemble methods such as Bagging. In bootstrap resampling, observations are drawn with replacement, which naturally leads to variability in the number of unique data points represented in each resampled dataset. Characterizing this distribution provides insight into the expected diversity of the bootstrap samples and, consequently, the degree of heterogeneity among the models in the ensemble. Moreover, in settings with class imbalance, the number of distinct observations from each class that appear in a bootstrap sample can deviate substantially from their proportions in the original dataset. Understanding the distribution of distinct outcomes therefore helps quantify how bootstrap sampling may amplify or mitigate imbalance across resamples. This perspective clarifies both the statistical properties of Bagging and the mechanisms through which bootstrap diversity contributes to model robustness.

Example 4.1 (Expected Representation of Weighted Minority Classes). *Consider a categorical random variable N taking values in $\{1, 2, 3\}$ with the imbalanced probability distribution*

$$P(N = 1) = p_1 = 0.80, \quad P(N = 2) = p_2 = 0.15, \quad P(N = 3) = p_3 = 0.05.$$

Here, class 1 is the majority class, while classes 2 and 3 represent minority groups with different weights. Suppose we draw $n = 3$ observations of N with replacement, and let X denote the number of distinct classes observed in these three draws. By Theorem 3.6, the expected number of distinct outcomes is

$$E(X) = \sum_{k=1}^3 (1 - (1 - p_k)^3) = 1 - (1 - p_1)^3 + 1 - (1 - p_2)^3 + 1 - (1 - p_3)^3.$$

We compute each term:

$$\begin{aligned} 1 - (1 - p_1)^3 &= 1 - (0.2)^3 = 0.992, \\ 1 - (1 - p_2)^3 &= 1 - (0.85)^3 \approx 0.386, \\ 1 - (1 - p_3)^3 &= 1 - (0.95)^3 \approx 0.143. \end{aligned}$$

Thus,

$$E(X) \approx 0.992 + 0.385875 + 0.142625 = 1.5205.$$

This calculation shows that, in expectation, a bootstrap sample of size 3 contains only about 1.52 distinct classes. The majority class appears with probability 0.992, meaning it is almost always represented. The heavier minority class ($p_2 = 0.15$) appears with probability 0.386, while the rarest minority class ($p_3 = 0.05$) appears with probability only 0.143. The theorem therefore quantifies how weighted minority classes contribute to the expected diversity of bootstrap samples and highlights how rare classes may be systematically underrepresented in small resamples.

4.2. Impact of Lottery Weighting on Pre-Lottery Position Diversity. Before turning to the NBA data, it is useful to note that the number of distinct outcomes in a weighted lottery provides a natural measure of how widely the lottery spreads opportunity among participants. In a system where each participant (or team) is assigned a probability weight, the distribution of distinct outcomes captures the likelihood that multiple participants appear among the winning positions, rather than a small subset repeatedly occupying them. This perspective is especially relevant in settings where the weighting scheme is intentionally designed to balance fairness, unpredictability, and competitive incentives.

The following example examines the NBA Draft Lottery in the 2018 and 2019 seasons, illustrating how teams occupy distinct pre-lottery positions under the league’s weighted lottery structure and how the distribution of these positions changes across the two years due to modifications in the lottery rules.

Example 4.2 (Distinct pre-lottery positions in the NBA). *Using Remark 3.1, Theorem 3.4, and Theorem 3.6, we compute the probabilities and expected values for the number of distinct pre-lottery positions observed in the NBA Draft Lottery for the 2018 and 2019 seasons:*

	2018	2019
$P[X = 1]$	3.880×10^{-9}	3.583×10^{-12}
$P[X = 6]$	0.3214	0.1171
$P[X = 7]$	0.3096	0.2826
$P[X = 8]$	0.1461	0.3280
$P[X = 14]$	1.783×10^{-10}	3.662×10^{-8}
$E(X)$	6.5059	7.7297

TABLE 4.2. Probabilities and expected values for the number of distinct pre-lottery positions that receive the top draft pick over 14 years using both 2018 and 2019 odds.

We conclude that the NBA’s changes to the draft lottery increased the expected number of pre-lottery positions receiving the top pick over a 14-year period from 6.5059 to 7.7297, representing an increase of approximately 19%. This finding is supported by the following data in Figure 1 garnered by sampling a 14-year period of the NBA Draft Lottery 10^5 times using both 2018 and 2019 lottery probabilities.

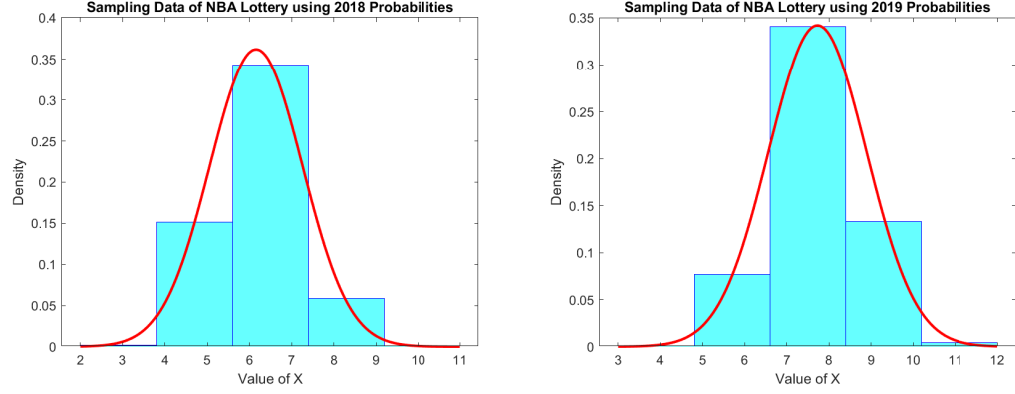


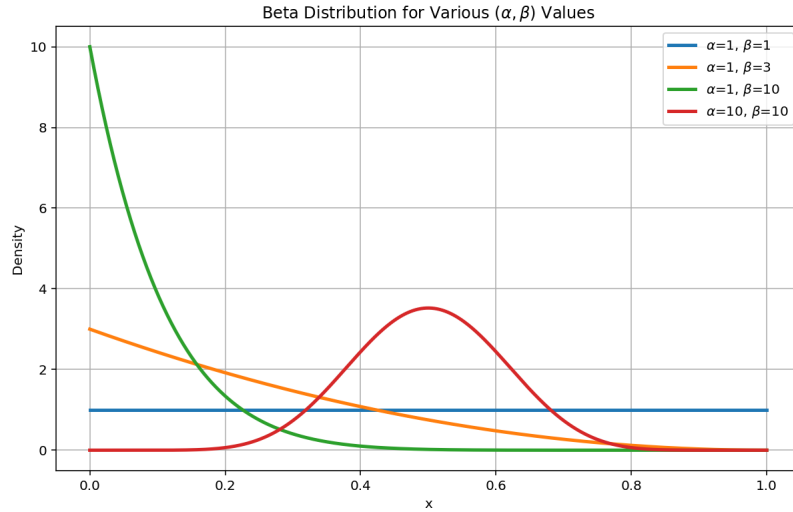
FIGURE 1. Simulations of the NBA Draft Lottery using 2018 and 2019 probabilities.

5. MONTE CARLO SIMULATIONS

We strive to gain a better understanding of both weighted and unweighted experiments by utilizing Monte Carlo simulations (MCS), a powerful computational tool. Here, our MCS utilize a beta random variable B with the density function

$$B(\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}.$$

Note that the beta random variable can represent both a weighted and uniform continuous distribution over $(0, 1)$.

FIGURE 2. Beta distribution for various (α, β) values.

To generate a discrete weighted distribution across n categories indexed from 1 to n , we map a continuous Beta distribution into discrete intervals using the floor function. Let $Y \sim \text{Beta}(\alpha, \beta)$ be a continuous random variable bounded on the interval $(0, 1)$. We define the discrete random variable $T(\alpha, \beta)$ as:

$$T(\alpha, \beta) = \lfloor n \cdot Y \rfloor + 1. \quad (5.28)$$

This transformation scales the continuous support by n , partitions it into n equal-width intervals via the floor operator $\lfloor \cdot \rfloor$, and shifts the resulting support from $[0, n-1]$ to $[1, n]$. By varying the shape parameters α and β , we alter the underlying probability density function of Y , which directly dictates the probability weights assigned to each discrete outcome. For instance, $(\alpha = 1, \beta = 1)$ yields a uniform discrete distribution; $(\alpha = 1, \beta = 3)$ and $(\alpha = 1, \beta = 10)$ introduce progressively stronger right-skewed profiles where lower-indexed categories receive higher weights; and $(\alpha = 10, \beta = 10)$ produces a symmetric, bell-shaped discrete distribution centered around the middle categories, as seen in Figure 2. Sampling from $T(\alpha, \beta)$ to generate a sequence of random numbers allows for a direct tally of distinct outcomes to analyze the empirical distribution under these varied weighting schemes. To produce sufficient data on this experiment, we executed this process m times and took the data's mean and variance at the end of the m -th experiment. Letting $n = 100$ and $m = 10^5$, we ran experiments (MCS) that varied the parameters α and β , which yielded the following data used to calculate the sample mean and sample variance for the trials.

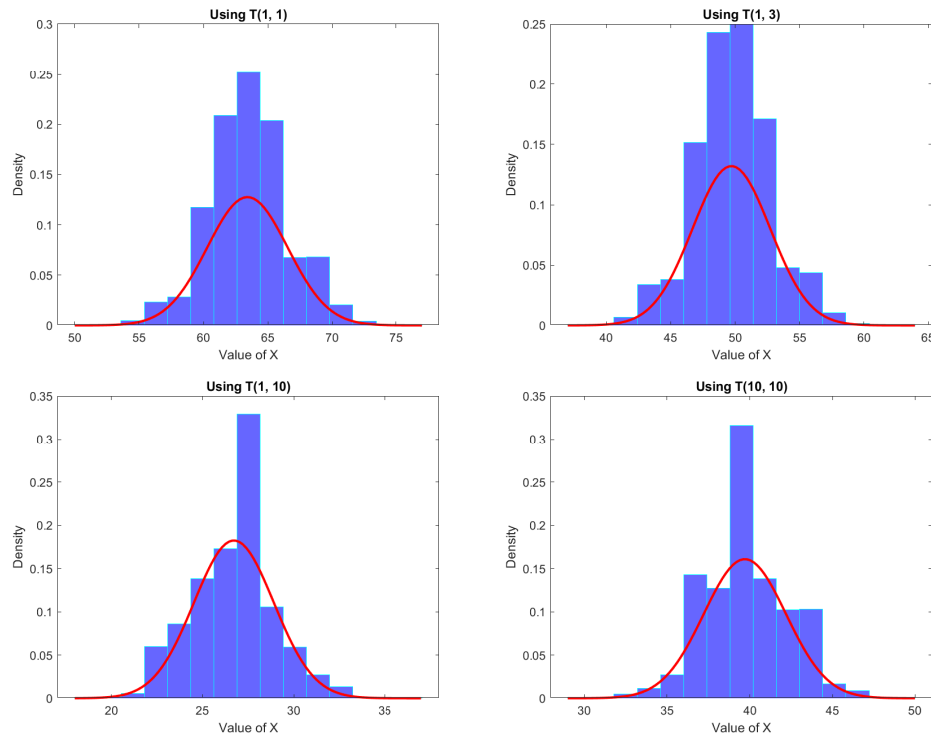


FIGURE 3. Histograms illustrating $m = 10^5$ MCS of X for $n = 100$ outcomes.

The mean and variance from the MCS are expressed in Table 5.3 below.

(α, β)	(1, 1)	(1, 3)	(1, 10)	(10, 10)
$\bar{X}$	63.4092	49.7069	26.7178	39.6961
$\text{Var}(X)$	9.7719	9.1289	4.7803	6.1499

TABLE 5.3. Mean and variance of $m = 10^5$ MCS for α and β with $n = 10^5$.

From the results of our MCS, we observe that the mean of our experiment is maximized when $\alpha = \beta = 1$, which represents the uniform distribution on $(0, 1)$ where each $p_i = 1/n$.

6. RESPONSE TO THE CONJECTURE IN [5]

The authors of [5] employed similar MCS results to conjecture that the expected value of the random variable X , representing the number of distinct outcomes in repeated experiments, attains its maximum when the events are equiprobable. We address this question by establishing the following theorem.

Theorem 6.1. *Let $n \geq 2$. The expected value of the number of distinct outcomes observed in a repeated weighted experiment with event probabilities $p_1, p_2, \dots, p_n$ is maximized when the events are equiprobable, i.e.,*

$$p_1 = \dots = p_n = \frac{1}{n}.$$

Proof. First, note that the expected value $E(X) = \sum_{k=1}^n [1 - (1 - p_k)^n] := \mathcal{E}(p_1, \dots, p_n)$ is concave down as its Hessian matrix $H(p_1, \dots, p_n) = \nabla^2 \mathcal{E}(p_1, \dots, p_n)$ is strictly negative definite with entries

$$h_{ii} = -n(n-1)(1-p_i)^{n-2} < 0 \quad \text{and} \quad h_{ij} = 0 \quad \text{when } i \neq j.$$

Therefore, $\mathcal{E}(p_1, \dots, p_n)$ has a maximizer. Now, we employ the Lagrange multiplier method to maximize $\mathcal{E}(p_1, \dots, p_n)$ subject to the probability constraint $\sum_{k=1}^n p_k = 1$. To this end, consider the function

$$F(p_1, \dots, p_n, \lambda) = \sum_{k=1}^n [1 - (1 - p_k)^n] + \lambda \left(1 - \sum_{k=1}^n p_k \right).$$

Observe that for each $1 \leq i \leq n$, we have

$$\frac{\partial F}{\partial p_i} = n(1 - p_i)^{n-1} - \lambda.$$

Setting each $\frac{\partial F}{\partial p_i} = 0$ we obtain

$$\lambda = n(1 - p_1)^{n-1} = \dots = n(1 - p_n)^{n-1},$$

implying that $p_1 = \dots = p_n$. Using $\sum_{k=1}^n p_k = 1$, we get $p_1 = \dots = p_n = 1/n$, as desired. $\square$

For a simple example, we use Figure 4 to illustrate that the expected value of two events is maximized at the point $(1/2, 1/2)$. This is done by considering $E(X)$ as a three-dimensional surface where $x = p_1$, $y = p_2$, and $z = E(X)$. Applying our expected value formula in Theorem 3.6 yields the surface $z = 2 - (1 - x)^2 - (1 - y)^2$, which we maximize subject to the constraint that $x + y = 1$ to produce Figure 4. Note that the probabilities in Figure 4 are represented by the parabolic intersection of the expected value surface and its constraining plane.

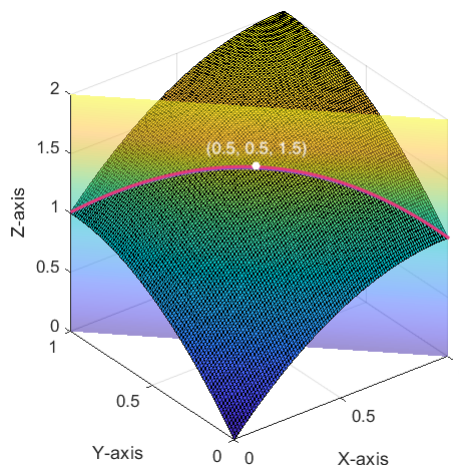


FIGURE 4. The expected value $z = 2 - (1 - x)^2 - (1 - y)^2$ is maximized at $(1/2, 1/2)$ subject to $x + y = 1$, where $x = p_1$ and $y = p_2$.

Remark 6.1. *The connection between our results and information theory is best understood through the framework developed by Kendall [11], who traces the entropy concept from Boltzmann’s statistical mechanics through Shannon’s foundational communication theory. Shannon’s measure of entropy,*

$$H = -K \sum_{i=1}^S p_i \log p_i,$$

is maximized when all outcomes are equiprobable. Shannon characterized H axiomatically through three requirements: continuity in the p_i , monotone increase in S under equiprobability (i.e., when $p_i = 1/S$), and an additivity postulate ensuring that H decomposes consistently across successive selections. Our Theorem 6.1 establishes an analogous maximization result in the discrete probabilistic setting of repeated weighted experiments: the expected number of distinct outcomes

$$E(X) = \sum_{k=1}^n [1 - (1 - p_k)^n]$$

is maximized when $p_1 = \dots = p_n = 1/n$. This result is consistent with the maximum-entropy principle, since the uniform distribution is precisely the distribution that maximizes Shannon entropy over a finite alphabet. Kendall emphasizes that entropy arises naturally as an additive quantity in thermodynamic and communication-theoretic contexts, and our expected value formula reflects a parallel additivity: $E(X)$ decomposes as a sum of contributions $1 - (1 - p_k)^n$ from each outcome, each representing the probability that outcome k appears at least once in n trials. Just as concentrating probability mass on fewer outcomes reduces Shannon entropy below its uniform maximum, concentrating probability weights reduces the expected diversity of observed outcomes below the equiprobable benchmark, a phenomenon our Monte Carlo simulations in Section 5 illustrate concretely across several Beta-distributed weighting schemes.

7. FURTHER APPLICATIONS

These games fit our model as each outcome in poker and craps has a well-defined probability, which increases our model's accuracy. To conclude, our findings can also apply to a different set of situations where probabilities are approximated using a sample set of data or a statistical model. This can be used to estimate the number of distinct seeds that win a knockout tournament [13], the number of distinct days that contain the winning ticket of a sweepstakes giveaway [3], and the number of distinct selection positions that win "The Singled Out Game" [19]. Furthermore, we believe our models can be used to analyze new variations of classic problems such as "The Birthday Problem" [10, 15], the problem of playing a fair game of craps [16] with a loaded pair of dice, and the De Moivre probability problem on the sum of numerical results in equally likely trials [8]. The endless possibilities to which our findings can be applied make them a versatile result that will provide great mathematical insights for future generations.

Acknowledgments. The authors would like to thank the referees for some useful comments and their helpful suggestions that have improved the quality of this paper.

Conflict of interest. The authors declare no conflict of interest.

REFERENCES

- [1] Ayanoglu, M. B., & Uysal, I. (2025). PWFS: Probability-weighted feature selection. *Electronics*, 14(11), Article 2264.
- [2] Cardano, G. (2015). *The book on games of chance: The 16th-century treatise on probability* (S. H. Gould & S. S. Wilks, Trans.). Dover Publications.
- [3] Ecker, M. W. (2007). Maximizing the probability of a big sweepstakes win. *The College Mathematics Journal*, 38(1), 32–36.
- [4] Edmonds, C. (2022). *What is tanking? A history of the NBA's controversial lottery strategy*.
- [5] Foster, D., & Yolcu, T. (2025). On the probability distribution of the number of distinct outcomes in repeated trials. *Mathematics Magazine*, 98(2), 96–109.
- [6] Friedgut, E. (2004). Hypergraphs, entropy, and inequalities. *The American Mathematical Monthly*, 111(9), 749–760.
- [7] Huygens, C. (1714). *Libellus de ratiociniis in ludo aleae: The value of all chances in games of fortune; cards, dice, wagers, lotteries, &c. mathematically demonstrated*. Printed by S. Keimer, for T. Woodward.
- [8] Jackson, C. S. (1917). A problem in probability. *The American Mathematical Monthly*, 24(2), 73–76.
- [9] Jiang, X., Zhou, X., & Kosorok, M. R. (2024). Deep doubly robust outcome weighted learning. *Machine Learning*, 113, 815–842.
- [10] Joag-Dev, K., & Proschan, F. (1992). Birthday problem with unlike probabilities. *The American Mathematical Monthly*, 99(1), 10–12.
- [11] Kendall, M. G. (1973). Entropy, probability and information. *International Statistical Review*, 41(1), 59–68.
- [12] Kruppa, J., Liu, Y., Biau, G., Kohler, M., König, I. R., Malley, J. D., & Ziegler, A. (2014). Probability estimation with machine learning methods for dichotomous and multicategory outcome. *Biometrical Journal*, 56(4), 534–563.
- [13] Kulhanek, T., & Ponomarenko, V. (2020). Surprises in knockout tournaments. *Mathematics Magazine*, 93(3), 193–199.
- [14] Lightner, J. E. (1991). A brief look at the history of probability and statistics. *The Mathematics Teacher*, 84(8), 623–630.

- [15] McKinney, E. H. (1966). Generalized birthday problem. *The American Mathematical Monthly*, 73(4), 385–387.
- [16] Morrison, I., & Swinarski, D. (2016). Can you play a fair game of craps with a loaded pair of dice? *The American Mathematical Monthly*, 123(2), 136–148.
- [17] NBA.com Staff. (2024). *2024 NBA draft lottery: Odds, history, and how it works*.
- [18] RealGM, L.L.C. (2025). *NBA draft lottery results*.
- [19] Shelton, K. (2005). The singled out game. *Mathematics Magazine*, 78(1), 15–25.

(C. Brinkman) SIEBEL SCHOOL OF COMPUTING AND DATA SCIENCE, UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN, CHAMPAIGN, IL 61820, USA

(L. Xue) DEPARTMENT OF MATHEMATICS, BRADLEY UNIVERSITY, 1501 W BRADLEY AVE, PEORIA, 61625, IL, USA

(T. Yolcu) DEPARTMENT OF MATHEMATICS, BRADLEY UNIVERSITY, 1501 W BRADLEY AVE, PEORIA, 61625, IL, USA