# WHEN ROOKS MISS: PROBABILITY THROUGH CHESS

STEVEN J. MILLER, HAOYU SHENG, AND DANIEL TUREK

ABSTRACT. A famous (and hard) chess problem asks what is the maximum number of safe squares possible in placing $n$ queens on an $n \times n$ board. We examine related problems from placing $n$ rooks. We prove that as $n \to \infty$, the probability rapidly tends to 1 that the fraction of safe squares from a random placement converges to $1/e^2$. Our interest in the problem is showing how to view the involved algebra to obtain the simple, closed form limiting fraction. In particular, we see the power of many of the key concepts in probability: binary indicator variables, linearity of expectation, variances and covariances, Chebyshev's inequality, and Stirling's formula.

## CONTENTS

## 1. INTRODUCTION

Probability as a subject owes a lot to the theory of games; many of its early results were inspired by attempts to analyze the odds of various configurations occurring. What began in contests involving dice or cards has grown tremendously over the centuries, and almost any subject can be the source of a good problem. In this note we concentrate on an extension of a chess question. There is no dearth of great sources for interesting chess problems (see for example [Pet1, Pet2, Wat]); we study one of them, the $n$ queens problem, with a twist. The only familiarity with chess we need is that queens move (and hence attack) horizontally, vertically and diagonally while rooks move just horizontally and vertically, and instead of the standard $8 \times 8$ board ours is $n \times n$. We assume the reader is familiar with the basics of probability and combinatorics, specifically that $\binom{n}{k} = n!/k!(n-k)!$ is the number of ways to choose $k$ objects from $n$ when order does not matter.

The standard definition of the problem is the following: *If we place $n$ queens on an $n \times n$ chessboard, what is the maximum number of pawns that can be placed and not attacked by any queen?* The problem is often tweaked to asking not just for the maximum number of pawns, but how many configurations of $n$ queens realize this. We can immediately convert this to probability questions by asking what is the probability a random configuration leaves the maximum number of squares safe, or what is the expected number of safe squares. While it is trivial to write down an approach to solve the first problem (it can be encoded as a simple linear programming problem), the run-time is tremendous and the optimal number of safe squares is known only for small $n$; starting with $n = 1$, the known values are

$$0, \ 0, \ 0, \ 1, \ 3, \ 5, \ 7, \ 11, \ 18, \ 22, \ 30, \ 36, \ 47, \ 56, \ 72, \ 82, \ 97, \ 111,$$
$$132, \ 145, \ 170, \ 186, \ 216, \ 240, \ 260, \ 290, \ 324, \ 360, \ 381, \ 420, \ldots \qquad (1.1)$$

though for $n \geq 17$ the values stated are merely probably the correct ones (see [LV, OEIS]. We can convert these to the percentage of the board that is safe in annnnnnnnnnnnnnnnnnnnnnn optimal configuration; the percentage starts at zero and is slowly increasing, reaching about 50% by the time $n = 30$; it can trivially be shown to converge to 100% in the limit.[1]

We study a variant which not only turns out to be easy to analyze, but also provides an excellent opportunity to see in action many important concepts in probability, including binary indicator variables, expected values, covariances, Chebyshev's inequality, Stirling's formula ($n! = n^n e^{-n} \sqrt{2\pi n} \, (1 + O(1/n))$), partitioning probabilities, and how to handle algebra to obtain clean limiting answers. (Recall big-Oh notation $f(x) = O(g(x))$ means that there is a $C > 0$ such that $|f(x)| \leq Cg(x)$.) For more on these and related topics, see for example [Mi].

***Rook Problem:*** *If $n$ rooks are randomly and uniformly placed on an $n \times n$ chessboard, what is the probability that at least one square is safe, what is the number of safe squares, and what is the distribution of the number of safe squares?*

It is of course possible that no squares are safe. If we have an $n \times n$ board and want to place $n$ queens, they could attack every square even if there are both rows and columns with no queens. For example, if $n = 4$ place the four queens in the corners; no queens are in rows 2 or 3, or columns 2 or 3, but every square is attacked. The situation is very different for rooks, as rooks can only attack horizontally and vertically, while queens also attack diagonally.

For there to be no safe squares, it must be the case that either each row has a rook, or each column has a rook.[2] Each rook is in a different column with probability $n^n/(n!\binom{n^2}{n})$. To see this, we argue by labeling the rooks 1 through $n$, so order matters. There are $n$ choices for the row of each rook, giving us $n^n$ configurations that have a rook in each column. There are $n^2$ squares to

---

[1]Essentially place the $n$ queens in a $\sqrt{n} \times \sqrt{n}$ square in the bottom right corner, which shows that only on the order of $n^{3/2}$ of the $n^2$ squares are attacked. If instead of queens we place rooks, it is not too hard to show that this is optimal; see Problem #1368 (by Steven J. Miller and Chenyang Sun) of the Spring 2020 issue of the Pi Mu Epsilon Journal.

[2]If there are no rooks in the $i^{\text{th}}$ column and none in the $j^{\text{th}}$ row, then the square $(i, j)$ is safe. We can bound the probability of a safe square by doubling this probability; to find the exact probability we would then subtract the probability that each row *and* each column has a rook.

place the first rook, $n^2 - 1$ for the second, and so on down to $n^2 - (n-1)$ for the $n^{\text{th}}$ rook; the product of this is $n!\binom{n^2}{n}$. As $n!\binom{n^2}{n} \leq (n^2)^n$, we have

$$\frac{n^n}{n!\binom{n^2}{n}} \leq \frac{n^n}{n^{2n}} = \frac{1}{n^n}, \tag{1.2}$$

which rapidly goes to zero with $n$. Thus the configurations where each row or each column has a rook happen with vanishingly small probability. In particular, this means that with probability quickly approaching 1 we have at least one square safe.

Thus the first part of our question is not too interesting; the second, however, is more involved and it is here that good mathematics and statistics enter.

**Theorem 1.1.** *As $n \to \infty$, the expected number of safe squares in the Rook Problem converges to $n^2/e^2$, so the percentage of safe squares converges to approximately 13.53%.*

The idea of the proof is the following. Let $\mathcal{B}$ be a board configuration (i.e., it is the location of $n$ squares on an $n \times n$ board where the rooks are placed), and consider the binary indicator variables

$$X_{ij}(\mathcal{B}) = \begin{cases} 1 & \text{if square } (i, j) \text{ is safe under } \mathcal{B} \\ 0 & \text{otherwise.} \end{cases} \tag{1.3}$$

Each configuration $\mathcal{B}$ occurs with probability $1/\binom{n^2}{n}$, as we must choose $n$ of the $n^2$ squares for our $n$ rooks.

We compute

$$\nu_n = \sum_{i,j=1}^{n} \mathbb{E}[X_{ij}(\mathcal{B})], \quad \sigma_n^2 = \sum_{ij=1}^{n} \mathbb{E}[X_{ij}(\mathcal{B}) - \nu_n)^2]; \tag{1.4}$$

note $\nu_n/n^2$ is the average fraction of safe squares. We then use Chebyshev's inequality to show that the number of safe squares for a random $\mathcal{B}$ is concentrated tightly near the mean, answering our second question. Unfortunately the $X_{ij}$ are not independent, and thus we cannot use the variance of a sum is the sum of the variance. Instead, we expand by using covariances.

In §2 we isolate and prove a general combinatorial lemma on ratios of binomial coefficients which is used in calculating the mean and the standard deviation. Doing so allows us to do the computation once, and then we just need to change the values of the parameters for the two cases. We go through the argument in complete detail; this is a major motivation of this paper, as one of our goals is to provide commentary on a very difficult part of mathematics: how does one do the algebra to obtain a nice, simple closed form expression? Such simplifications are not always possible, but when they are there is tremendous value, as we now have exact answers and do not need to resort to simulations. After this, in §3 we prove our main result about the limiting percentage of safe squares. We also look at our main results using simulation.

## 2. COMBINATORIAL PRELIMINARIES

Our key ingredient in the analysis is the following combinatorial lemma involving ratios of binomial coefficients, which we use to calculate the probability of certain rook configurations. We

have $n^2$ squares initially available for the rooks, and will see later as we go through all the cases that there are always integers $a$ and $b$ such that the number of squares we lose is $an + b$.

**Lemma 2.1.** *For any positive integer $a$ and any integer $b$,*

$$\lim_{n\to\infty} \frac{\binom{n^2-an-b}{n}}{\binom{n^2}{n}} = \frac{1}{e^a}. \tag{2.1}$$

*Proof.* We have

$$
\begin{aligned}
\frac{\binom{n^2-an-b}{n}}{\binom{n^2}{n}} &= \frac{(n^2-an-b)!}{n!(n^2-(a+1)n-b)!} \cdot \frac{n!(n^2-n)!}{(n^2)!} \\
&= \frac{(n^2-an-b)!}{(n^2-(a+1)n-b)!} \cdot \frac{(n^2-n)!}{(n^2)!} \\
&= \frac{(n^2-n)(n^2-n-1)\cdots(n^2-(a+1)n-(b-1))}{n^2(n^2-1)\cdots(n^2-an-(b-1))} \\
&= \prod_{k=0}^{an+b-1} \frac{n^2-n-k}{n^2-k} = \prod_{k=0}^{an+b-1} \left(1 - \frac{n}{n^2-k}\right).
\end{aligned} \tag{2.2}
$$

The challenge now is to manipulate the product so that we can obtain a nice limit as $n \to \infty$. Notice we have on the order of $n$ terms, each within a fixed multiple of $1/n$ from 1. Thus we should be thinking that the answer is related to the exponential function, as

$$e^x = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n. \tag{2.3}$$

This *suggests* how we should attack the algebra; we want to view each term $1 - n/(n^2 - k)$ as $1 - 1/n$ plus a *very* small correction, a correction significantly smaller than $1/n$ and so small that it will not contribute as $n$ tends to infinity, even when raised to the $n^{\text{th}}$ power. Once we realize this, the following expansion is forced upon us:

$$\frac{n}{n^2-k} = \frac{1}{n} + \frac{k}{n^3-nk}. \tag{2.4}$$

We can often get a good feeling for an expression by looking at extreme cases. Since $k$ is between 0 and $an + b - 1$, we have

$$\left(1 - \frac{1}{n} - \frac{an+b-1}{n^3-an^2-(b-1)n}\right)^{an+b} \leq \prod_{k=0}^{an+b-1}\left(1 - \frac{n}{n^2-k}\right) \leq \left(1 - \frac{1}{n}\right)^{an+b}. \tag{2.5}$$

Since $a$ and $b$ are fixed integers, the far right is easily analyzed:

$$\lim_{n\to\infty}\left(1 - \frac{1}{n}\right)^{an+b} = \left(\lim_{n\to\infty}\left(1 - \frac{1}{n}\right)^n\right)^a \cdot \left(\lim_{n\to\infty}\left(1 - \frac{1}{n}\right)\right)^b = e^a. \tag{2.6}$$

Similar to above, we have a sense of the size of the factor on the far left: $1 - 1/n$. We pull this factor out, as its behavior to the power $an + b$ is the same as our analysis of the upper bound. If we can just show the remaining factor is extremely close to 1 then the upper and lower bounds for our

product will be close, and we can then take limits as $n \to \infty$. Since

$$
\begin{aligned}
1 - \frac{1}{n} - \frac{an+b-1}{n^3 - an^2 - (b-1)n} &= \left(1 - \frac{1}{n}\right)\left(1 - \frac{an+b-1}{n^3 - an^2 - (b-1)n} \cdot \frac{n}{n-1}\right) \\
&= \left(1 - \frac{1}{n}\right)\left(1 - \frac{an+b-1}{n^2 - an - (b-1)} \cdot \frac{1}{n-1}\right),
\end{aligned} \tag{2.7}
$$

substituting this into the left of (2.5) yields

$$
\left(1 - \frac{1}{n}\right)^{an+b}\left(1 - \frac{an+b-1}{n^2 - an - (b-1)} \cdot \frac{1}{n-1}\right)^{an+b} \leq \prod_{k=0}^{an+b-1}\left(1 - \frac{n}{n^2 - k}\right) \leq \left(1 - \frac{1}{n}\right)^{an+b}. \tag{2.8}
$$

We only need to find the limit for $\left(1 - \frac{an+b-1}{n^2-an-(b-1)} \cdot \frac{1}{n-1}\right)^{an}$ as $n \to \infty$ to find the limit for the left. As we only care about its limit as $n \to \infty$, we can be a bit crude in approximations in finding upper and lower bounds, so long as they have the same limit.

A good upper bound is trivial:

$$
\lim_{n\to\infty}\left(1 - \frac{an+b-1}{n^2 - an - (b-1)} \cdot \frac{1}{n-1}\right)^{an} \leq \lim_{n\to\infty} 1^{an} = 1. \tag{2.9}
$$

We thus want a lower bound that tends to 1 as $n \to \infty$. For large $n$ the factor is essentially $1 - a/n^2$, and thus when we raise this to the $n^{\text{th}}$ power it should tend to 1 (from the definition of the exponential function, we would need to raise it to something on the order of $n^2$ to have a limit that is not 1). Thus we can be very crude in our bounding.[3] Fix any $c > 0$. Since $a$ is a positive integer, as $n \to \infty$ we have $\frac{an+b-1}{n^2-an-(b-1)} \leq \frac{c}{n}$. Thus

$$
\lim_{n\to\infty}\left(1 - \frac{an+b-1}{n^2 - an - (b-1)} \cdot \frac{1}{n-1}\right)^{an} > \lim_{n\to\infty}\left(1 - \frac{c}{n}\right)^{an} = \left(\lim_{n\to\infty}\left(1 - \frac{c}{n}\right)^{n}\right)^{a} = e^{-ca}. \tag{2.10}
$$

As this is true for any $c > 0$, taking the limit as $c$ tends to zero from above we obtain a lower bound of $e^0 = 1$. Since the upper and lower bounds are equal, by the Squeeze Theorem we obtain

$$
\lim_{n\to\infty}\left(1 - \frac{an+b-1}{n^2 - an - (b-1)} \cdot \frac{1}{n-1}\right)^{an} = 1, \tag{2.11}
$$

and therefore

$$
\lim_{n\to\infty}\frac{\binom{n^2-an-b}{n}}{\binom{n^2}{n}} = \lim_{n\to\infty}\prod_{k=0}^{an+b-1}\left(1 - \frac{n}{n^2 - k}\right) = \frac{1}{e^a}, \tag{2.12}
$$

completing the proof. $\qquad \square$

## 3. PROOF OF LIMITING PERCENTAGE

Let $S_n(\mathcal{B})$ be the number of safe spaces on an $n \times n$ board when the $n$ rooks are in configuration $\mathcal{B}$ and thus $S_n(\mathcal{B})/n^2$ is the fraction of squares that are safe; therefore

$$
S_n(\mathcal{B}) = \sum_{i,j=1}^{n} X_{ij}(\mathcal{B}). \tag{3.1}
$$

---

[3]We could be more careful, and take logarithms, analyze and then exponentiate, but there is no need.

As the $X_{ij}$ are Bernoulli random variables, to calculate their expected values we just need to know the probability it takes on the value 1 (or the value 0). For a square $(i, j)$ to be safe we cannot place a rook in row $i$ or in column $j$; there are $2n - 1$ squares in the union of row $i$ and column $j$, and thus the $n$ rooks must be placed in the $n^2 - (2n - 1) = (n - 1)^2$ remaining spaces on the board. Thus the probability that $(i, j)$ is safe is $\binom{(n-1)^2}{n} / \binom{n^2}{n}$, and therefore

$$\mathbb{E}[X_{ij}] \;=\; 1 \cdot \mathrm{Prob}(X_{ij} = 1) + 0 \cdot \mathrm{Prob}(X_{ij} = 0) \;=\; \frac{\binom{(n-1)^2}{n}}{\binom{n^2}{n}} \;=:\; \mu_n. \tag{3.2}$$

### 3.1. Determining the Mean.

The mean of $S_n$ is easily determined (and from that we can get the percentage of safe squares by dividing by $n^2$). By linearity of expectation, and all the random variables being identically distributed,

$$\mathbb{E}[S_n] \;=\; \sum_{i,j=1}^{n} \mathbb{E}[X_{ij}] \;=\; n^2 \mathbb{E}[X_{11}] \;=\; n^2 \frac{\binom{(n-1)^2}{n}}{\binom{n^2}{n}}, \tag{3.3}$$

and therefore the expected percentage of the $n \times n$ square that is safe is

$$\mathbb{E}[S_n/n^2] \;=\; \frac{\binom{(n-1)^2}{n}}{\binom{n^2}{n}} \;=\; \mu_n. \tag{3.4}$$

We now use Lemma 2.1; as $(n - 1)^2 = n^2 - 2n + 1$ we have $a = 2$ and $b = -1$, and thus

$$\mu \;:=\; \lim_{n \to \infty} \mu_n \;=\; \lim_{n \to \infty} \mathbb{E}[S_n/n^2] \;=\; \lim_{n \to \infty} \frac{\binom{n^2 - 2n + 1}{n}}{\binom{n^2}{n}} \;=\; \frac{1}{e^2} \;\approx\; 13.53\%. \tag{3.5}$$

It is remarkable that all the binomial ratios interact beautifully and a clean, simple closed form solution is available for the limiting percentage of safe squares when we randomly and uniformly place $n$ rooks on an $n \times n$ board: $\mu = 1/e^2$.

We study this limiting percentage through simulation, and indeed find this limiting percentage of approximately 13.53% (Figure 1).

### 3.2. Determining the Variance.

We now turn to the next part: how does the percentage of safe squares vary around $1/e^2$? To determine the answer, we need to compute the variance of $S_n/n^2$ or $S_n$:

$$\mathrm{Var}(S_n) \;=\; \sum_{i,j=1}^{n} \mathrm{Var}(X_{ij}) + \sum_{\substack{i,j,\ell,m=1 \\ (i,j) \neq (\ell,m)}}^{n} \mathrm{Cov}(X_{ij}, X_{\ell m}). \tag{3.6}$$

We analyze the two pieces separately.

#### 3.2.1. The Variance Term.

From the definition of variance,

$$\mathrm{Var}(X_{ij}) \;=\; \mathbb{E}[X_{ij}^2] - \mathbb{E}[X_{ij}]^2. \tag{3.7}$$

As the $X_{ij}$ are binary indicator variables, $\mathbb{E}[X_{ij}^2] = \mathbb{E}[X_{ij}]$ and thus

$$\mathrm{Var}(X_{ij}) \;=\; \mathbb{E}[X_{ij}] - \mathbb{E}[X_{ij}]^2 \;=\; \mu_n - \mu_n^2. \tag{3.8}$$
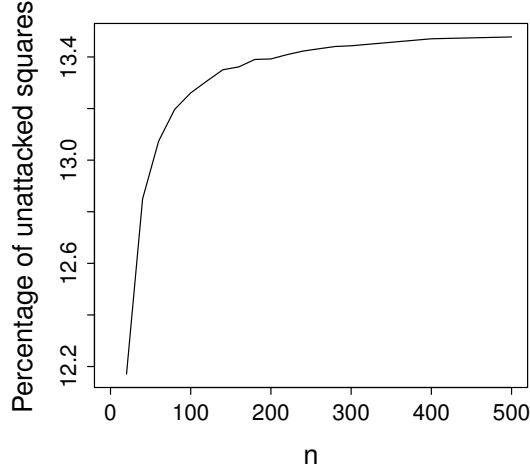
6

FIGURE 1. Mean percentage of unattacked squares, on an $n \times n$ board.

Thus the contribution of these terms to the variance of $S_n/n^2$ is

$$\frac{1}{n^2} \sum_{i,j=1}^{n} \text{Var}(X_{ij}) \;=\; \frac{1}{n^2} \cdot n^2(\mu_n - \mu_n^2) \;=\; \mu_n - \mu_n^2, \tag{3.9}$$

which in the limit as $n \to \infty$ is just

$$\mu - \mu^2 \;=\; \frac{1}{e^2} - \frac{1}{e^4}. \tag{3.10}$$

3.2.2. *The Covariance Term.* Breaking down the covariances, we have

$$\begin{aligned} \text{Cov}\,(X_{ij}, X_{\ell m}) \;&=\; \mathbb{E}[(X_{ij} - \mu_{ij})\,(X_{\ell m} - \mu_{\ell m})] \\ &=\; \mathbb{E}[X_{ij}X_{\ell m}] - \mathbb{E}[\mu_{\ell m}X_{ij}] - \mathbb{E}[\mu_{ij}X_{\ell m}] + \mu_{ij}\mu_{\ell m}. \end{aligned} \tag{3.11}$$

As the random variables are identically distributed and all have expected value $\mu_n$,

$$\text{Cov}\,(X_{ij}, X_{\ell m}) \;=\; \mathbb{E}[X_{ij}X_{\ell m}] - \mu_n^2. \tag{3.12}$$

As our variables are binary, note $X_{ij}X_{\ell m}(b)$ can be understood as

$$X_{ij}X_{\ell m}(\mathcal{B}) \;=\; \begin{cases} 1 & \text{if position } (i,j) \text{ and position } (\ell, m) \text{ are safe in board configuration } \mathcal{B} \\ 0 & \text{otherwise.} \end{cases}$$
$$\tag{3.13}$$

There are thus two cases to consider for the calculation of $\mathbb{E}[X_{ij}X_{lm}]$.

7

**Lemma 3.1.** *Assume the pairs $(i, j)$ and $(\ell, m)$ share either the same row or the same column (they cannot share both as they are distinct points, since $(i, j) \neq (\ell, m)$). Then*

$$Prob\,(X_{ij}X_{lm} = 1) \;=\; \mathbb{E}[X_{ij}X_{\ell m}] \;=\; \frac{\binom{(n-1)(n-2)}{n}}{\binom{n^2}{n}}, \tag{3.14}$$

*and the number of such pairs is $2(n-1)n^2$.*

*Proof.* If the two pairs share the same row, there are $\binom{n}{1}$ ways to choose the row and $\binom{n}{2}$ ways to choose the column. There are 2! ways to assign the columns to $(i, j)$ and $(\ell, m)$, and thus there are $2 \cdot n \cdot n(n-1)/2 = n^2(n-1)$ such pairs. There are the same number of pairs that share a column but not a row, and thus the number of pairs in this case is $2(n-1)n^2$.

What is the probability that $X_{ij}X_{\ell m}(\mathcal{B})$ equals 1 for such a pair? For convenience, let's assume they share a row (so $i = \ell$). We need to place $n$ rooks, none can be in row $i$, and none can be in column $j$ or $m$. We thus have $(n-1) \cdot (n-2)$ squares on the board where we may place the rooks, and the probability all $n$ rooks are in these squares is $\binom{(n-1)(n-2)}{n}/\binom{n^2}{n}$. Thus

$$\mathbb{E}[X_{ij}X_{\ell m}] \;=\; \frac{\binom{(n-1)(n-2)}{n}}{\binom{n^2}{n}}, \tag{3.15}$$

as claimed, and the contribution from all these pairs is

$$2(n-1)n^2 \frac{\binom{(n-1)(n-2)}{n}}{\binom{n^2}{n}} \;=\; 2(n-1)n^2 \frac{\binom{n^2-3n+2}{n}}{\binom{n^2}{n}}. \tag{3.16}$$

$\square$

**Lemma 3.2.** *Assume the pairs $(i, j)$ and $(\ell, m)$ are in distinct rows and columns (so $i \neq \ell$ and $j \neq m$). Then*

$$Prob\,(X_{ij}X_{lm} = 1) \;=\; \mathbb{E}[X_{ij}X_{\ell m}] \;=\; \frac{\binom{(n-2)^2}{n}}{\binom{n^2}{n}}, \tag{3.17}$$

*and the number of such pairs is $n^2(n-1)^2$.*

*Proof.* There are $n^2$ ways to choose $(i, j)$, and then $(n-1)^2$ choices for $(\ell, m)$ so that it is not in the same row or column as $(i, j)$. What is the probability that $X_{ij}X_{\ell m}(\mathcal{B})$ equals 1 for such a pair? We need to place $n$ rooks, none can be in row $i$ or $\ell$, and none can be in column $j$ or $m$. We thus have $(n-2) \cdot (n-2)$ squares on the board where we may place the rooks, and the probability all $n$ rooks are in these squares is $\binom{(n-2)(n-2)}{n}/\binom{n^2}{n}$. Thus

$$\mathbb{E}[X_{ij}X_{\ell m}] \;=\; \frac{\binom{(n-2)(n-2)}{n}}{\binom{n^2}{n}}, \tag{3.18}$$

as claimed, and the contribution from all these pairs is

$$n^2(n-1)^2 \frac{\binom{(n-2)(n-2)}{n}}{\binom{n^2}{n}} \;=\; n^2(n-1)^2 \frac{\binom{n^2-4n+3}{n}}{\binom{n^2}{n}}. \tag{3.19}$$

$\square$

**Remark 3.3.** *As a consistency check, note the number of pairs from Lemmas 3.1 and 3.2 is*

$$2(n-1)n^2 + n^2(n-1)^2 \;=\; n^2\left((2n-2) + (n^2-2n+1)\right) \;=\; n^2(n^2-1); \qquad (3.20)$$

*as expected, this equals the number of distinct pairs $(i,j)$ and $(\ell,m)$.*

We can now determine the contribution of the covariance terms in (3.6):

$$\sum_{\substack{i,j,\ell,m=1 \\ (i,j)\neq(\ell,m)}}^{n} \mathrm{Cov}\,(X_{ij}, X_{\ell m}) \;=\; \sum_{\substack{i,j,\ell,m=1 \\ (i,j)\neq(\ell,m)}}^{n} \left(\mathbb{E}[X_{ij}\mathbb{E}_{lm}] - \mu_n^2\right)$$

$$= \; 2(n-1)n^2 \frac{\binom{(n-1)(n-2)}{n}}{\binom{n^2}{n}} + n^2(n-1)^2 \frac{\binom{(n-2)^2}{n}}{\binom{n^2}{n}} - n^2 \cdot (n^2-1)\mu_n^2. \tag{3.21}$$

### 3.3. **Variance of $S_n$.** Using the results from the previous subsections, we can write $\mathrm{Var}(S_n)$ as

$$\mathrm{Var}(S_n) \;=\; \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Var}(X_{ij}) + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{n} \sum_{m=1, i,j\neq l,m}^{n} \mathrm{Cov}(X_{ij}, X_{lm})$$

$$= \; n^2(\mu_n - \mu_n^2) + 2(n-1)n^2 \frac{\binom{(n-1)(n-2)}{n}}{\binom{n^2}{n}} + n^2(n-1)^2 \frac{\binom{(n-2)^2}{n}}{\binom{n^2}{n}} - n^2 \cdot (n^2-1)\mu_n^2$$

$$= \; n^2(\mu_n - \mu_n^2) + (2n^3 - 2n^2) \frac{\binom{(n-1)(n-2)}{n}}{\binom{n^2}{n}} + (n^4 - 2n^3 + n^2) \frac{\binom{(n-2)^2}{n}}{\binom{n^2}{n}} - (n^4 - n^2)\mu_n^2 \tag{3.22}$$

As $\mathrm{Var}(S_n/n^2) = \mathrm{Var}(S_n)/n^4$, if we divide both sides by $n^4$ we obtain

$$\mathrm{Var}(S_n/n^2)$$

$$= \; \frac{\mu_n - \mu_n^2}{n^2} + \left(\frac{2}{n} - \frac{2}{n^2}\right) \frac{\binom{n^2-3n+2}{n}}{\binom{n^2}{n}} + \left(1 - \frac{2}{n} + \frac{1}{n^2}\right) \frac{\binom{n^2-4n+4}{n}}{\binom{n^2}{n}} - \left(1 - \frac{1}{n^2}\right)\mu_n^2. \tag{3.23}$$

We do not need to determine the variance perfectly, only bound it well enough so we can apply Chebyshev's inequality. We use Lemma 2.1 to evaluate the two ratios of Binomial coefficients; in the first we take $a = 3$ and $b = -4$, for a limit of the ratio of $1/e^3$; for the second term we have $a = 4, b = b = -4$ and get $1/e^2$. The product $\mu_n^2$ converges to $1/e^2$.

Substituting, we see that the constant term (in $n$) above is zero due to cancelation, and the coefficient of the $1/n$ term, as $n \to \infty$, converges to $2/e^3 - 2/e^4$. As the remaining terms are of size $1/n^2$, we see that for large $n$

$$\mathrm{Var}(S_n/n^2) \;\approx\; \frac{2}{e^3}\left(1 - \frac{1}{e}\right)\frac{2}{n}; \tag{3.24}$$

in particular, for any $\epsilon > 0$ for all $n$ sufficiently large we have

$$\sigma_n^2 \;:=\; \mathrm{Var}(S_n/n^2) \;\leq\; \frac{(2+\epsilon)(e-1)}{e^3}\frac{1}{n} \;:=\; \frac{C_\epsilon}{n}. \tag{3.25}$$

9

By Chebyshev's inequality,

$$\text{Prob}\left(\left|\frac{S_n}{n^2} - \mu_n\right| \geq k\sigma_n\right) \leq \frac{1}{k^2}. \tag{3.26}$$

If we take $k = \log n$ then the probability above converges to zero, and $k\sigma_n \leq C_\epsilon \log n/n$ also converges to zero. This means for any fixed $\epsilon' > 0$ that as $n \to \infty$ the probability of $S_n/n^2$ being more than $\epsilon'$ from $\mu_n$ tends to zero; as $\mu_n \to \mu = 1/e^2$, we see that in the limit for almost all boards the percentage of safe squares can be made to be as close to $1/e^2$ as we wish.

## 4. Future Work

The result from §3.3 is an example of the power of Chebyshev's inequality. Simply by bounding the variance we are able to show that $S_n/n^2$ concentrates on the mean; unfortunately it is not strong enough to say anything about the fluctuations about the mean.

We again turn to simulation, to examine the nature of these fluctuations about the mean. Simulations suggest that the distribution of free squares is asymptotically Gaussian (Figure 2), however we do not attempt to prove this result.
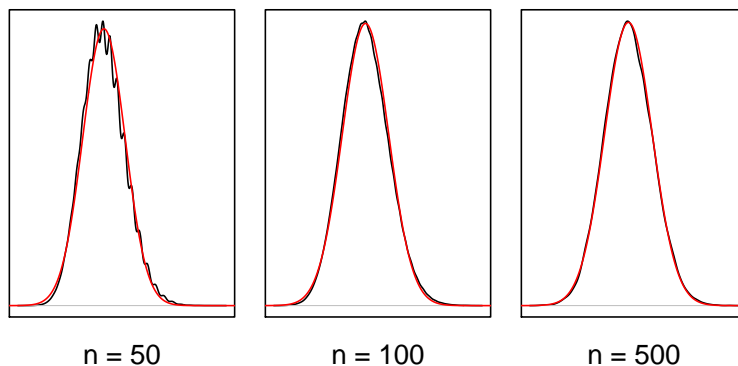


$$n = 50 \qquad\qquad n = 100 \qquad\qquad n = 500$$

FIGURE 2. Distribution of the percentage of free squares for various board sizes $n$. Empirical distribution from 10,000 simulations (black), Gaussian density of same mean and variance (red).

There are many other problems one could investigate. For example, what if instead of queens and rooks one were to use bishops? Is it possible to obtain similar results, and if so what combinatorial estimates are now needed?

## References

[LV]    B. Lemaire and P. Vitushinkiy, *Placing $n$ non dominating queens on the $n \times n$ chessboard, Part I*, French Federation of Mathematical Games, 2011. https://www.ffjm.org/upload/fichiers/N_NON_DOMINATING_QUEENS.pdf.
[Mi]    S. J. Miller, *The Probability Lifesaver*, Princeton University Press, Princeton, NJ, 2018. https://web.williams.edu/Mathematics/sjmiller/public_html/probabilitylifesaver/index.htm.
[OEIS]  OEIS, *Sequence A001366*, https://oeis.org/A001366.
[Pet1]  M. Petkovic, *Mathematics and chess : 110 entertaining problems and solutions*, Dover Publications, 1997.
[Pet2]  M. Petkovic, *Mathematics and Chess*, Dover Recreational Math, 2011.
[Wat]   J. J. Watkins, *Across the Board: The Mathematics of Chessboard Problems*, Princeton University Press, 2004.

*Email address*: sjm1@williams.edu,Steven.Miller.MC.96@aya.yale.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

*Email address*: hs9@williams.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

*Email address*: dbt1@williams.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267