

# **PROBABILITY THEORY II**

Matthew Joseph, notes by Ramdas Singh

Second Semester

# List of Symbols

$\Omega$ , a sample space.

$\omega$ , an element of a sample space.

$EX$ , the expectation of the random variable  $X$ .

$\text{Var}X$ , the variance of the random variable  $X$ .

$N(\mu, \sigma^2)$ , a normal distribution with expectation  $\mu$  and variance  $\sigma^2$ .

# Contents

1	1
1.1 The Law of Large Numbers . . . . .	2
1.2 Simple Random Walk . . . . .	3
1.3 Erdős-Renyi Random Graph . . . . .	7
Appendices	11
A Appendix	13
Index	15

# Chapter 1

January 3rd.

We first start with some initial statements. Let  $\Omega$  be a countable state space, and let each  $\omega \in \Omega$  have a probability  $P(\omega)$  associated with it.

**Lemma 1.1.** *For random variables  $X, Y$  such that  $X(\omega) \leq Y(\omega)$  for all  $\omega \in \Omega$ . Then,  $EX \leq EY$ .*

*Proof.* This can easily be seen by summing over all terms via the alternate definition of the expectation,

$$EX = \sum_{\omega \in \Omega} X(\omega)P(\omega) \leq \sum_{\omega \in \Omega} Y(\omega)P(\omega) = EY. \quad (1.1)$$

■

We now state Markov's inequality.

**Theorem 1.2** (*Markov's inequality*). *If  $X$  is a non-negative random variable, then for  $a > 0$ , we have*

$$P(X > a) \leq \frac{EX}{a}. \quad (1.2)$$

*Proof.* Define an indicator function  $I_a(\omega)$  as 1 if  $X(\omega) \geq a$ , and 0 if otherwise. We then have

$$I_a(\omega) \leq \frac{X(\omega)}{a} \implies P(X \geq a) = EI_a \leq \frac{1}{a}EX. \quad (1.3)$$

■

**Remark 1.3.** A better upper bound here may be found by starting with  $I_a(\omega)X(\omega)$  instead of just  $X(\omega)$ .

If we have  $X \sim N(0, 1)$ , then we can find an upper bound for its probability density function.

$$P(X > a) = \int_a^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \leq \int_a^\infty \frac{1}{\sqrt{2\pi}} \frac{x}{a} e^{-\frac{x^2}{2}} dx = \frac{e^{-\frac{a^2}{2}}}{\sqrt{2\pi}a}. \quad (1.4)$$

Note that  $X$  here is a random variable over a continuous state space; the previous lemma and Markov's inequality also work here. We are to show them for the continuous case instead of the discrete one.

*Proof.* Here, we have  $0 \leq X(\omega) \leq Y(\omega)$  for all  $\omega$  in our continuous state space  $\Omega$ . We see that  $\{X > x\} \subseteq \{Y > x\} \implies P(X > x) \leq P(Y > x)$ . Integrating both sides gives us  $EX \leq EY$ . ■

**Theorem 1.4** (*Chebyshev's inequality*). *Let  $X$  be a random variable with finite mean  $\mu = EX$  and finite variance  $\sigma^2 = \text{Var}(X)$ . Then for  $a > 0$ ,*

$$P(|X - \mu| > a) \leq \frac{\text{Var}(X)}{a^2}. \quad (1.5)$$

*Proof.* Start with the proof of Markov's inequality, replacing the indicator function with one that's unity when  $|X - \mu| \geq a$ . ■

**Example 1.5.** Suppose  $X_1, X_2, \dots, X_n$  are  $n$  independent and identically distributed random variables, with  $EX_i = \mu$  and  $\text{Var}X_i = \sigma^2$ . If  $S_n = \sum X_i$ , we then have

$$P(|S_n - n\mu| > a) \leq \frac{\text{Var}S_n}{a^2} = \frac{n\sigma^2}{a^2}. \quad (1.6)$$

If we replace  $a$  with  $n^{\frac{1}{2}+\varepsilon}$ , we then have

$$P(|S_n - n\mu| > n^{\frac{1}{2}+\varepsilon}) \leq \frac{\sigma^2}{n^{2\varepsilon}} \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (1.7)$$

**Proposition 1.6.** If  $\text{Var}(X) = 0$ , then  $P(X = EX) = 1$ .

*Proof.* For all  $\varepsilon > 0$ , we have

$$P(|X - EX| > \varepsilon) \leq \frac{\text{Var}X}{\varepsilon^2} = 0. \quad (1.8)$$

Define  $A_n$  as  $\{|X - EX| > \frac{1}{n}\}$ . Taking  $P(\bigcup A_n) = \lim_{n \rightarrow \infty} P(A_n)$ , the proof follows. ■

## 1.1 The Law of Large Numbers

We start by stating the weak law of large numbers.

**Theorem 1.7** (*Weak law of large numbers*). Let  $\{X_k\}_{k \geq 1}$  be a sequence of independent and identically distributed random variables with  $E|X_i| < \infty$ . Let  $\mu = EX_i$ . Then for any  $a > 0$ ,

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - \mu\right| > a\right) = 0. \quad (1.9)$$

*Proof.* For now, let us assume that  $\Omega$  is countable. We begin with the case where the variance of  $X_i$ ,  $\sigma^2$ , is finite. Fix  $a > 0$ , and let  $S_n = X_1 + X_2 + \dots + X_n$ . Then,

$$P\left(\left|\frac{S_n}{n} - \mu\right| > a\right) = P(|S_n - n\mu| > na) \leq \frac{\text{Var}S_n}{n^2a^2} = \frac{n\sigma^2}{n^2a^2} \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (1.10)$$

We now focus the case when the variance,  $\sigma^2$ , is infinite. Assume that the expected value,  $\mu$ , is 0; if it were non-zero, we would then instead work with  $X_i - \mu$ . Let  $\delta > 0$ ; we shall choose a particular  $\delta$  later. For each  $n$ , define  $n$  pairs of random variables,  $U_1, V_1, \dots, U_n, V_n$ , as  $U_k = X_k, V_k = 0$  if  $|X_k| \leq \delta n$ , and  $U_k = 0, V_k = X_k$  if  $|X_k| > \delta n$ .  $X_k$  can be rewritten as  $U_k + V_k$ . We then have

$$\{|X_1 + \dots + X_n| \geq na\} \subseteq \{|U_1 + \dots + U_n| \geq \frac{na}{2}\} \cup \{|V_1 + \dots + V_n| \geq \frac{na}{2}\} \quad (1.11)$$

$$\implies P(|X_1 + \dots + X_n| \geq na) \leq P\left(|U_1 + \dots + U_n| \geq \frac{na}{2}\right) + P\left(|V_1 + \dots + V_n| \geq \frac{na}{2}\right). \quad (1.12)$$

We focus on the first term on the right hand side. The  $U_i$ 's are independently and identically distributed, so

$$P\left(|U_1 + \dots + U_n| \geq \frac{na}{2}\right) \leq \frac{4E[|U_1 + \dots + U_n|^2]}{a^2n^2} = \frac{4}{a^2n^2} (\text{Var}(U_1 + \dots + U_n) + (nEU_i)^2). \quad (1.13)$$

For the variance, we have

$$\text{Var}(U_1 + \dots + U_n) = n\text{Var}U_i \leq nEU_i^2 \leq nE[|U_i||U_i|] \leq \delta n^2 E[|U_i|] \quad (1.14)$$

which transforms the previous equation as

$$P\left(|U_1 + \dots + U_n| \geq \frac{na}{2}\right) \leq \frac{4}{a^2n^2} (\delta n^2 E[|U_i|] + (nEU_i)^2). \quad (1.15)$$

A lemma (to be proven later) states that  $E[|U_i|] = E[|X_i|]$  as  $n \rightarrow \infty$ , and  $EU_i = EX_i = 0$  too. So,

$$P\left(|U_1 + \dots + U_n| \geq \frac{na}{2}\right) \leq \frac{4}{a^2n^2} (\delta n^2 E[|U_i|] + (nEU_i)^2) \leq \frac{4\delta E[|U_i|]}{a^2} + \frac{4}{a^2} (EU_i)^2. \quad (1.16)$$

For the second term on the right hand side, begin with

$$\begin{aligned} P(V_1 + \dots + V_n \neq 0) &\leq P(\{V_1 \neq 0\} \cup \dots \cup \{V_n \neq 0\}) \leq nP(V_i \neq 0) = n \sum_{|x| > \delta n} P(X_i = x) \\ &\leq n \sum_{|x| > \delta n} \frac{|x|}{\delta n} P(X_i = x) = \frac{1}{\delta} E[|V_i|]. \end{aligned} \quad (1.17)$$

The rightmost term here tends to 0 as  $n \rightarrow \infty$ . Now choose  $\delta$  to be  $\frac{\varepsilon a^2}{6E[|X_i|]}$ , and then choose  $N$  to be large enough such that for all  $n > N$ , both the terms are smaller than  $\frac{\varepsilon}{2}$ . ■

*January 7th.*

We now prove the lemma called upon earlier.

**Lemma 1.8.** *If  $X$  is a discrete random variable and takes values  $y_1, y_2, \dots, y_k$ , and  $E[|X|] < \infty$ , then  $\lim_{n \rightarrow \infty} E[|X| \cdot 1_{|X| \leq n}] = E[|X|]$ .*

*Proof.* Notice that the terms on the left hand side and right hand side are  $\sum_{y_k: |y_k| \leq n}$  and  $\sum_{y_k} |y_k| P(Y = y_k)$ . The condition for convergence may now be applied. ■

The above equation, begin inside absolute braces, must imply that the term  $E[X \cdot 1_{|X| \leq n}]$  must also absolutely converge to  $EX$ .

## 1.2 Simple Random Walk

Let  $X_1, X_2, \dots$  be independent and identically distributed random variables, with  $X_i = 1$  with probability  $\frac{1}{2}$  and  $X_i = -1$  with probability  $\frac{1}{2}$ . Now define  $S_0 = 0$  and  $S_n = \sum_{i=1}^n X_i$ . The sequence  $(S_n)_{n \geq 0}$  is a *simple random walk*.

Note that  $S_0 = k_0 = 0, S_1 = k_1, \dots, S_n = k_n$  can occur if and only if  $|k_i - k_{i+1}| = 1$  for all  $0 \leq i \leq n-1$ . The sequence  $(k_n)_{n \geq 0}$  is a *simple path* of the simple random walk. By the event  $\{S_n = k\}$ , we are concerned with the event that the random walk visits  $k$  at step  $n$ . If  $(k_n)_{n \geq 0}$  is given we have  $X_i = k_i - k_{i-1}$ . Because the  $X_i$ 's are independent and identically distributed, each event  $\{X_1 = l_1, X_2 = l_2, \dots, X_n = l_n\}$ , where  $l_i = \pm 1$ , is equally likely with probability  $\frac{1}{2^n}$ . Thus,

$$P(S_n = k) = \frac{N_n(k)}{2^n} \quad (1.18)$$

where  $N_n(k)$  is defined as the number of distinct of path that start at 0 and end at  $k$  at step  $n$ . We also define  $N_n^+(k)$  to be the number of distinct paths that end at  $k$  at step  $n$  and stay above the  $x$ -axis up to time  $n-1$ . The probability of the corresponding event is

$$P(\{S_1 > 0, S_2 > 0, \dots, S_{n-1} > 0, S_n = k\}) = \frac{N_n^+(k)}{2^n}. \quad (1.19)$$

**Lemma 1.9.** *Suppose  $a, a', b, b'$  are integers, with  $0 \leq a < a'$ . Then the number of distinct path from  $(a, b)$  to  $(a', b')$  depends only on  $a' - a = n$  and  $b' - b = k$ , and is given by  $\binom{n+k}{\frac{n+k}{2}}$ .*

*Proof.* Notice that we need  $x+1$ 's and  $y-1$ 's to appear, satisfying  $x+y = a' - a$  and  $x-y = b' - b$ . Solving, we get  $x = \frac{n+k}{2}$  and  $y = \frac{n-k}{2}$ . Thus, the number of paths is given by  $\binom{n+k}{\frac{n+k}{2}}$ . ■

Using this lemma, we find that  $N_n(k) = \binom{n+k}{\frac{n+k}{2}}$ . The following convention is now followed; if  $t$  is not an integer, then  $\binom{n}{t} = 0$ .

**Lemma 1.10** (The *method of images*). *Suppose  $a, a', b, b'$  are integers, with  $0 \leq a < a'$  and  $b, b' > 0$ . Then the number of distinct paths from  $(a, b)$  to  $(a', b')$  that intersect the  $x$ -axis is equal to the number of paths from  $(a, -b)$  to  $(a', b')$ .*

*Proof.* Consider any path  $(b = k_0, k_1, \dots, k_{n-1}, k_n = b')$ , from  $(a, b)$  to  $(a', b')$ , that intersects the  $x$ -axis. Let  $j$  be the smallest index for which  $k_j = 0$ . For ease, denote  $(a, b)$  by  $A$ ,  $(a', b')$  by  $A'$ ,  $(a+j, 0)$  by  $B$ , and  $(a, -b)$  by  $A''$ . Reflect the segment from  $A$  to  $B$  about the  $x$ -axis to obtain a 'mirrored-path' from  $A''$  to  $B$ ;  $(-b = -k_0, -k_1, \dots, -k_{j-1}, k_j = 0, k_{j+1}, \dots, k_n = b')$ . There is now a one-to-one correspondence between the paths from  $A$  to  $A'$  that intersect the  $x$ -axis, and the paths from  $A''$  to  $A'$ . ■

We can now easily compute  $N_n^+(k)$ ; it simply the number of paths from  $(1, 1)$  to  $(n, k)$  that do not intersect the  $x$ -axis.

**Theorem 1.11** (*Ballot theorem*). *The number of paths that progress from  $(0, 0)$  to  $(n, k)$  through strictly positive values is given by  $N_n^+(k) = \frac{k}{n} N_n(k)$ .*

*Proof.* We have

$$\begin{aligned} N_n^+(k) &= \text{number of paths from } (1, 1) \text{ to } (n, k) - \text{number of such paths that intersect the } x\text{-axis} \\ &= N_{n-1}(k-1) - N_{n-1}(k+1) \\ &= \binom{n-1}{\frac{n+k}{2}-1} - \binom{n-1}{\frac{n+k}{2}} = \frac{k}{n} \binom{n}{\frac{n+k}{2}} = \frac{k}{n} N_n(k). \end{aligned} \quad (1.20)$$

■

Suppose  $n = 2\nu$ . Define  $u_{2\nu}$  to be  $P(S_{2\nu} = 0) = \frac{\binom{2\nu}{\nu}}{2^{2\nu}}$ . The question we ask is to compute the probability that the first return to 0, if at all, occurs after step  $n$ . It can be found out as

$$P(\text{first return to } 0 \dots) = P(S_1 \neq 0, S_2 \neq 0, \dots, S_{2\nu} \neq 0) \quad (1.21)$$

$$= P(S_1 > 0, \dots, S_{2\nu} > 0) + P(S_1 < 0, \dots, S_{2\nu} < 0)$$

$$= 2P(S_1 > 0, \dots, S_{2\nu} > 0)$$

$$= 2 \sum_{k \text{ even}, k > 0} P(S_1 > 0, \dots, S_{2\nu-1} > 0, S_{2\nu} = k)$$

$$= \frac{2}{2^{2\nu}} \sum_{k \text{ even}, k > 0} N_{2\nu}^+(k)$$

$$= \frac{2}{2^{2\nu}} \sum_{k \text{ even}, k > 0} N_{2\nu-1}(k-1) - N_{2\nu-1}(k+1)$$

$$= \frac{2}{2^{2\nu}} N_{2\nu-1}(1) = u_{2\nu}. \quad (1.22)$$

We state this down as a lemma.

**Lemma 1.12** (*Basic lemma*). *For  $n$  even, the probability that the first return to 0, if at all, occurs after step  $n$  is the same as the probability that the location at step  $n$  is 0. For  $n$  odd, it is the probability that the location at step  $n-1$  is 0.*

We ask another question; for a fixed  $n$ , where does the random walk achieve its first maximum upto time  $n$ ? For this, denote by  $M_n$  the index  $m$  at which the walk  $S_0, S_1, \dots, S_n$ , over  $n$  steps, achieves its maximum for the first time.

For  $0 < m < n$ ,  $M_n = m$  if and only if  $S_m > S_0, S_m > S_1, \dots, S_m > S_{m-1}$  and  $S_m \geq S_{m+1}, S_m \geq S_{m+2}, \dots, S_m \geq S_n$ . Notice that the first of these two conditions depends only on  $X_1, X_2, \dots, X_m$ , and the second condition depends only on  $X_{m+1}, X_{m+2}, \dots, X_n$ . So,  $P(M_n = m) = P(S_m > S_0, S_m > S_1, \dots, S_m > S_{m-1}) \cdot P(S_m \geq S_{m+1}, S_m \geq S_{m+2}, \dots, S_m \geq S_n)$ .

The key idea here is to consider the *reversed walk*; define a new walk with  $X'_1 = X_m, X'_2 = X_{m-1}, \dots, X'_m = X_1$ . Also define  $S'_k = X'_1 + \dots + X'_k$ . From here, we can deduce that  $S_m > S_{m-i}$  is true if and only if  $X_m + \dots + X_{m-i}$  is true, which is true if and only if  $S'_i > 0$  is true. So,  $P(S_m > S_0, S_m > S_1, \dots, S_m > S_{m-1}) = P(S'_1 > 0, S'_2 > 0, \dots, S'_m > 0)$ . If we now define  $S''_k = X_{m+1} + \dots + X_{m+k}$ , we have

$$\begin{aligned} P(S_m \geq S_{m+1}, S_m \geq S_{m+2}, \dots, S_m \geq S_n) &= P(X_{m+1} \leq 0, X_{m+1} + X_{m+2} \leq 0, \dots, X_{m+1} + \dots + X_n \leq 0) \\ &= P(S''_1 \leq 0, S''_2 \leq 0, \dots, S''_{n-m} \leq 0) \\ &= P(S''_1 \geq 0, S''_2 \geq 0, \dots, S''_{n-m} \geq 0) \end{aligned}$$

The first of the terms discussed,  $P(S'_1 > 0, S'_2 > 0, \dots, S'_m > 0)$ , can be computed for  $m = 2\nu, 2\nu + 1$ ; it is simply  $\frac{1}{2} u_{2\nu}$ . For the latter of these terms, we introduce a new random variable  $\tilde{X}$  which has the same distribution as the  $X_i$ 's and is independent. Also define  $\tilde{S}_i$  to be  $\tilde{X} + X_1 + \dots + X_{i-1}$  and  $\tilde{S}_0$  to be 0.

We then have

$$\begin{aligned}
\frac{1}{2}P(S_0 \geq 0, \dots, S_{n-m} \geq 0) &= P(\tilde{X} = 1) \cdot P(S_0 \geq 0, \dots, S_{n-m} \geq 0) \\
&= P(\tilde{X} = 1, S_0 \geq 0, S_0 \geq 0, \dots, S_{n-m} \geq 0) \\
&= P(\tilde{S}_1 = 1, \tilde{S}_2 > 0, \dots, S_{n-m+1} > 0) \\
&= P(S_1 > 0, S_2 > 0, \dots, S_{n-m+1} > 0).
\end{aligned} \tag{1.23}$$

Thus, we get

$$P(M_n = m) = \frac{1}{2}u_{2k}u_{2\nu-2k} \tag{1.24}$$

where  $m$  is of the form  $2k$  or  $2k + 1$ , and  $n$  is of the form  $2\nu$ , with  $1 < k < \nu$ .

*January 10th.*

Plugging in  $m = 0$ , we get  $P(M_n = 0) = P(S_1 \leq 0, \dots, S_{2\nu} \leq 0) = \frac{1}{2}u_{2\nu}$ . For  $m = n$ , we have  $P(M_n = n) = P(S_1 \leq 0, \dots, S_{2\nu} \leq 0) = \frac{1}{2}u_{2\nu}$ . Let us first compute  $u_{2k}$ .

$$\begin{aligned}
u_{2k} = P(2k = 0) &= \frac{\binom{2k}{k}}{2^{2k}} = \frac{(2k)!}{(k!)^2 2^{2k}} \\
&\sim \frac{(2k)^{2k+\frac{1}{2}} e^{-2k} \sqrt{2\pi}}{(\sqrt{2\pi} k^{k+\frac{1}{2}} e^{-k})^2 2^{2k}} = \frac{1}{\sqrt{\pi k}}.
\end{aligned} \tag{1.25}$$

For  $0 < a < b < 1$ , we have

$$\begin{aligned}
P(an \leq M_n \leq bn) &= \sum_{m=an}^{bn} P(M_n = m) = \sum_{k=a\nu}^{b\nu} u_{2k}u_{2\nu-2k} \\
&\sim \sum_{k=a\nu}^{b\nu} \frac{1}{\sqrt{\pi k}} \frac{1}{\sqrt{\pi(\nu-k)}} = \sum_{k=a\nu}^{b\nu} \frac{1}{\nu \sqrt{\pi \frac{k}{\nu}} \sqrt{\pi(1-\frac{k}{\nu})}} \\
&\rightarrow \frac{1}{\pi} \int_a^b \frac{dx}{\sqrt{x(1-x)}} = \frac{2}{\pi} (\arcsin \sqrt{b} - \arcsin \sqrt{a}).
\end{aligned} \tag{1.26}$$

In fact, this is the *arcsin law for maxima*; for  $0 \leq t \leq 1$ , we have

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n}{n} \leq t\right) = \frac{2}{\pi} \arcsin \sqrt{t}. \tag{1.27}$$

If we look at this as a cumulative density function, the probability density function becomes  $\frac{d}{dt} \frac{2}{\pi} \arcsin \sqrt{t} = \frac{1}{\pi \sqrt{t(1-t)}}$ .

We are now interested in  $\tilde{M}_n$ , the last time when maximum up to time  $n$  is attained. We can just look at the walk backwards again; in this case, we get

$$P\left(\frac{\tilde{M}_n}{n} \leq t\right) = P\left(\frac{n - \tilde{M}_n}{n} \leq t\right) \rightarrow \frac{2}{\pi} \arcsin \sqrt{t}. \tag{1.28}$$

We now ask the probability that the random walk of  $n = 2\nu$  steps last visit 0 at time  $2k$ . We denote by  $K_n$  the location of the last return to 0 in a walk of  $n$  steps. Now look at

$$\begin{aligned}
\alpha_{2k,2\nu} &= P(K_n = 2k) = P(S_{2k} = 0, S_{2k+1} \neq 0, \dots, S_{2\nu} \neq 0) \\
&= P(S_{2k} = 0) \cdot P(X_{2k+1} \neq 0, \dots, X_{2k+1} + \dots + X_{2\nu} \neq 0) \\
&= P(S_{2k} = 0) \cdot P(S_1 \neq 0, \dots, S_{2\nu-2k} \neq 0) = u_{2k}u_{2\nu-2k}.
\end{aligned} \tag{1.29}$$

We can also state an *arcsin law for last visit* here; for  $0 < t < 1$

$$\lim_{n \rightarrow \infty} P(K_n \leq tn) = \frac{2}{\pi} \arcsin \sqrt{t}. \tag{1.30}$$



If we set the an additional limit that says  $t$  tends to 0, replacing  $t$  by an arbitrary  $\varepsilon > 0$ , we have

$$\lim_{n \rightarrow \infty} P(K_n = 0) = 0. \quad (1.31)$$

Given enough time, a simple random walk must return to 0.

Denote by  $f_{2n}$  the probability that the first return to 0 occurs at time  $2n$ .

$$\begin{aligned} f_{2n} &= P(S_1 \neq 0, \dots, S_{2n-1} \neq 0, S_{2n} = 0) \\ &= P(S_1 \neq 0, \dots, S_{2n-1} \neq 0) - P(S_1 \neq 0, \dots, S_{2n} \neq 0) \\ &= P(S_1 \neq 0, \dots, S_{2n-2} \neq 0) - P(S_1 \neq 0, \dots, S_{2n} \neq 0) \\ &= u_{2n-2} - u_{2n} = \frac{1}{2n-1} u_{2n}. \end{aligned} \quad (1.32)$$

**Lemma 1.13.** *With the usual notation,*

$$u_{2n} = f_2 u_{2n-2} + f_4 u_{2n-4} + \dots + f_{2n} u_0. \quad (1.33)$$

*Proof.* We have

$$\begin{aligned} P(S_{2n} = 0) &= \sum_{k=1}^n P(S_{2n} = 0, \text{ first return at } 2k) \\ &= \sum_{k=1}^n P(\text{first return at } 2k) \cdot P(S_{2n} = 0 \mid \text{first return at } 2k) \\ \implies P(S_n = 0) &= \sum_{k=1}^n f_{2k} u_{2n-2k}. \end{aligned} \quad (1.34)$$

■

**Theorem 1.14.** *The probability that in the time interval 0 to  $n = 2\nu$ , the random walk spends  $2k$  amount of time on the positive side and  $2\nu - 2k$  amount of time on the negative side is  $\alpha_{2k, 2\nu}$ .*

**Corollary 1.15.** *For  $0 < t < 1$ ,*

$$P(\text{random walk spends less than } tn \text{ time on positive side}) \rightarrow \frac{2}{\pi} \arcsin \sqrt{t}. \quad (1.35)$$

*Proof.* This is the proof of the theorem. We introduce  $b_{2k, 2\nu}$ ; it is defined as the probability that the random walk of length  $2\nu$  and  $2k$  sides above the  $x$ -axis. We need to show that  $b_{2k, 2\nu} = \alpha_{2k, 2\nu}$ . We have

$$b_{2\nu, 2\nu} = P(S_1 \geq 0, S_2 \geq 0, \dots, S_{2\nu} \geq 0) = u_{2\nu}, \quad (1.36)$$

$$b_{0, 2\nu} = P(S_1 \leq 0, \dots, S_{2\nu} \leq 0) = u_{2\nu}. \quad (1.37)$$

We are left to prove it for  $1 \leq k \leq \nu - 1$ . Assume that exactly  $2k$  out of  $2\nu$  time are spent above the  $x$ -axis, with  $1 \leq k \leq \nu - 1$ . Suppose first return to 0 occurs at time  $2r < 2\nu$ . We deal in cases.

- Case I:  $2r$  time units upto first return are on the positive side. Then,  $r \leq k \leq \nu - 1$ . The time from  $2r$  to  $2\nu$  has to be above the  $x$ -axis,  $2k - 2\nu$  time. The number of such paths is  $(\frac{1}{2} 2^{2r} f_{2r})(2^{2\nu-2r} b_{2k-2r, 2\nu-2r})$ .
- The  $2r$  time units upto the first return are on the negative side. The nubmer of such paths is  $(\frac{1}{2} 2^{2r} f_{2r})(2^{2\nu-2r} b_{2k, 2\nu-2r})$ . Also,  $\nu - r \geq k$ .

Thus, we have

$$b_{2k, 2\nu} = \frac{1}{2} \sum_{r=1}^k f_{2r} b_{2k-2r, 2\nu-2r} + \frac{1}{2} \sum_{r=1}^{\nu-k} f_{2r} b_{2k, 2\nu-2r}. \quad (1.38)$$

We now proceed with induction on  $\nu$ . We have already shown this for  $\nu = 1$ ; assume that this is true for  $\nu \leq V - 1$ . By induction,

$$\begin{aligned} b_{2k, 2V} &= \frac{1}{2} \sum_{r=1}^k f_{2r} \alpha_{2k-2r, 2V-2r} + \frac{1}{2} \sum_{r=1}^{V-k} f_{2r} \alpha_{2k, 2V-2r} \\ &= \frac{1}{2} u_{2V-2k} \sum_{r=1}^k f_{2r} u_{2k-2r} + \frac{1}{2} u_{2k} \sum_{r=1}^{V-k} f_{2r} u_{2V-2k-2r} \\ &= u_{2k} u_{2V-2k} = \alpha_{2k, 2\nu}. \end{aligned} \quad (1.39)$$

■

January 17th.

**Theorem 1.16** (*Weirstrass's polynomial approximation.*). Let  $f : [0, 1] \rightarrow \mathbb{R}$  be a continuous function. Then for every  $\varepsilon > 0$ , there is a polynomial  $P$ , dependent on  $f$  and  $\varepsilon$ , such that

$$|f(x) - P(x)| < \varepsilon \text{ for all } x \in [0, 1]. \quad (1.40)$$

**Remark 1.17.** Any continuous function  $f : [0, 1] \rightarrow \mathbb{R}$  is bounded and uniformly continuous. This fact will be useful in proving the previous theorem.

*Proof.* Start with  $X_1, X_2, \dots$  which are independent and identically distributed Bernoulli random variables,  $\text{Ber}(x)$ . Let  $S_n = X_1 + X_2 + \dots + X_n$ . From the weak law of large numbers, we know that  $\frac{S_n}{n}$  is approximately  $x$ . We can expect that  $f(x)$  will also be approximately  $f(\frac{S_n}{n})$ . We now have

$$\begin{aligned} f_n(x) &= Ef\left(\frac{S_n}{n}\right) = \sum_{j=0}^n f\left(\frac{j}{n}\right)P(S_n = j) \\ &= \sum_{j=0}^n f\left(\frac{j}{n}\right)\binom{n}{j}x^j(1-x)^{n-j}. \end{aligned} \quad (1.41)$$

This is now a polynomial; we wish to see how close this is to  $f$ . Define  $A_\delta$  to be  $\{j : |\frac{j}{n} - x| \leq \delta\}$

$$\begin{aligned} |f_n(x) - f(x)| &= \left| \sum_{j=0}^n \left( f\left(\frac{j}{n}\right) - f(x) \right) P(S_n = j) \right| \\ &= \left| \sum_{j \in A_\delta} \left( f\left(\frac{j}{n}\right) - f(x) \right) + \sum_{j \notin A_\delta} \left( f\left(\frac{j}{n}\right) - f(x) \right) \right| P(S_n = j) \\ &\leq \sum_{j \in A_\delta} \left| f\left(\frac{j}{n}\right) - f(x) \right| P(S_n = j) + \sum_{j \notin A_\delta} \left| f\left(\frac{j}{n}\right) - f(x) \right| P(S_n = j). \end{aligned} \quad (1.42)$$

We have two terms to deal with now. For the first term, choose  $\delta > 0$  such that  $|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$ ; this  $\delta$  can be chosen since  $f$  is uniformly continuous. Similarly, also choose  $M = \sup_{x \in [0, 1]} |f(x)|$ .  $M$  is finite since  $f$  is bounded. Thus, we have

$$\sum_{j \in A_\delta} \left| f\left(\frac{j}{n}\right) - f(x) \right| P(S_n = j) \leq \sum_{j \in A_\delta} \varepsilon P(S_n = j) \leq \varepsilon \quad (1.43)$$

and

$$\sum_{j \notin A_\delta} \left| f\left(\frac{j}{n}\right) - f(x) \right| P(S_n = j) \leq 2MP\left(\left|\frac{S_n}{n} - x\right| > \delta\right) \leq 2M \frac{\text{Var}(S_n)}{n^2 \delta^2} = \frac{2Mnx(1-x)}{n^2 \delta^2}. \quad (1.44)$$

Combining the two, and choosing  $n$  large enough, we have

$$|f_n(x) - f(x)| \leq \varepsilon + \frac{2Mx(1-x)}{n\delta^2} \leq \varepsilon + \frac{M}{2n\delta^2} \leq 2\varepsilon. \quad (1.45)$$

■

### 1.3 Erdős-Renyi Random Graph

We first discuss the setup; start with  $n$  vertices of an empty graph. For any pair of points  $(i, j)$ , with  $i \neq j$ , join these vertices with an edge with probability  $p$  independently for all such pairs. Such a graph is denoted by  $G_{n,p}$ .

A collection of three points  $S = \{i, j, k\}$  form a triangle if  $G_{n,p}$  has the edges  $\{i, j\}$ ,  $\{j, k\}$ , and  $\{i, k\}$ . We question the probability that such a graph has no formed triangles. Can we find  $p = p_n$  such that

triangles begin to appear at  $p_n$ ? Let  $S$  be any set of three vertices. Define  $X_S$  to be the indicator function; 1 if  $S$  forms a triangle, and 0 otherwise. We note that  $X_S \sim \text{Ber}(p^3)$ . We note that

$$EX_S = p^3, \text{Var}X_S = p^3(1 - p^3) \leq p^3.$$

Denote by  $N$  the number of triangles in the graph  $G_{n,p}$ . Clearly,

$$N = \sum_{S:|S|=3} X_S, \quad EN = \binom{n}{3} p^3 < n^3 p^3, \quad \text{Var}N = \sum_S \text{Var}X_S + \sum_S \sum_{T \neq S} \text{Cov}(X_S X_T) \leq n^3 p^3 + n^4 p^5$$

Also,  $P(N \geq 1) \leq EN < n^3 p^3$ . If  $p = p_n < \frac{1}{n}$ , then  $P(N \geq 1) \rightarrow 0$  as  $n \rightarrow \infty$ . We discuss this for  $p > \frac{1}{n}$ . We have

$$P(N = 0) \leq P(|N - EN| \geq EN) \leq \frac{\text{Var}N}{(EN)^2} \leq \frac{(n^3 p^3 + n^4 p^5)}{\frac{n^6 p^6}{100}} \leq \frac{100}{n^3 p^3} + \frac{100}{n(np)} \rightarrow 0. \quad (1.46)$$

We can state this as a theorem.

**Theorem 1.18.** *Consider  $G_{n,p_n}$ . Let  $E$  be the event that the graph is triangle free. We then have*

$$P(E) \rightarrow \begin{cases} 0 & \text{if } \frac{p_n}{\frac{1}{n}} \rightarrow \infty, \\ 1 & \text{if } \frac{p_n}{\frac{1}{n}} \rightarrow 0. \end{cases} \quad (1.47)$$

Now suppose that  $\frac{np_n}{\frac{1}{n}} C > 0$  as  $n \rightarrow \infty$ . Then we have

$$N \approx \text{Poisson}\left(\frac{C^3}{6}\right). \quad (1.48)$$

January 21st.

**Remark 1.19.** For this next ‘game’, we will think of  $X_i$ ’s as the winnings in game  $i$  and  $\mu$  to be the entrance fees for a game.

**Definition 1.20.** Suppose that  $X_1, X_2, \dots$  are independent, but not necessarily identically distributed. Let  $S_n = X_1 + \dots + X_n$ . We say a game with accumulated entrance fees  $\{\alpha_n, n \geq 1\}$  is fair if

$$P\left(\left|\frac{S_n}{\alpha_n} - 1\right| > \varepsilon\right) \rightarrow 0 \quad (1.49)$$

for all  $\varepsilon > 0$ .

Using this definition of ‘fair’, we look at an example.

**Example 1.21.** This is the St. Petersburg’s paradox. This is the game; toss a coin repeatedly until the first head is observed. If this head occurs at the  $k^{\text{th}}$  toss, the amount paid out is  $X = 2^k$ . Let us find a fair accumulated entrance fees. In this case,

$$EX = \sum_{k=1}^{\infty} \frac{1}{2^k} 2^k = \infty. \quad (1.50)$$

Suppose we play this game  $n$  times. We are to find a fair accumulated sum  $\{\alpha_n\}$  such that

$$P(|S_n - \alpha_n| > \varepsilon \alpha_n) \rightarrow 0. \quad (1.51)$$

To find this, we will define

$$\begin{aligned} U_j &= X_j 1_{\{X_j \leq a_n\}}, \\ V_j &= X_j 1_{\{X_j > a_n\}}. \end{aligned}$$

$a_n$  shall be determined later. Note that  $S_n = X_1 + \dots + X_n = U_1 + \dots + U_n + V_1 + \dots + V_n$ . Then,

$$P(|S_n - \alpha_n| > \varepsilon \alpha_n) \leq P(|U_1 + \dots + U_n - \alpha_n| > \frac{1}{2} \varepsilon \alpha_n) + P(|V_1 + \dots + V_n| > \frac{1}{2} \varepsilon \alpha_n). \quad (1.52)$$

We first bound the second term on the right hand side. We have

$$P(|V_1 + \dots + V_n| > \frac{1}{2}\varepsilon\alpha_n) \leq P(\bigcup_{i=1}^n \{V_i \neq 0\}) \leq nP(V_1 \neq 0) = nP(X_1 > a_n) \quad (1.53)$$

$$= \sum_{2^k > a_n} P(X = 2^k) \leq \frac{2n}{a_n}. \quad (1.54)$$

Thus, we will require that  $a_n \gg n$ . Also,

$$EU_1 = \sum_{k \leq \log_2 a_n} 2^k \cdot 2^{-k} = \lfloor \log_2 a_n \rfloor, \quad \text{Var} U_1 \leq E[U_1^2] = \sum_{k \leq \log_2 a_n} (2^k)^2 \cdot 2^{-k} = 2^{\lfloor \log_2 a_n \rfloor + 1} - 1 < 2a_n. \quad (1.55)$$

$\frac{1}{n}(U_1 + \dots + U_n) \approx EU_j = \lfloor \log_2 a_n \rfloor$ , so we should choose

$$\alpha_n = nEU_j = n \lfloor \log_2 a_n \rfloor. \quad (1.56)$$

This gives us

$$P(|U_1 + \dots + U_n - \alpha_n| > \frac{1}{2}\varepsilon\alpha_n) \leq \frac{n(2a_n)}{\frac{1}{4}\varepsilon^2\alpha_n^2}. \quad (1.57)$$

Thus, we have another condition where we require that  $\frac{na_n}{\alpha_n^2} \rightarrow 0$ . The conditions we require are

$$\frac{n}{a_n} \rightarrow 0 \text{ and } \frac{na_n}{n^2(\log_2 a_n)^2} \rightarrow 0.$$

The sequence  $\{a_n\}$  defined as  $a_n = n \log_2 n$  satisfies these properties. The sequence  $\alpha_n$  is thus

$$\alpha_n = n \log_2 a_n = n \log_2 n + n \log_2 \log_2 n. \quad (1.58)$$



# Appendices



## Chapter A

# Appendix

Extra content goes here.





# Index

arcsin law for last visit, 5  
arcsin law for maxima, 5

Ballot theorem, 4  
Basic lemma, 4

Chebyshev's inequality, 1

Markov's inequality, 1

method of images, 3

reversed walk, 4

simple path, 3  
simple random walk, 3

Weak law of large numbers, 2

Weirstrass's polynomial approximation, 7