IGL Project Timeline

Nathan Sell

Last Updated: December 1, 2019

1 Background

1.1 Probability

When we use mathematics to discuss the outcomes of random experiments, and their probabilities, we start by assigning a random variable numbers which represent the outcomes of some experiment. For example, if we let X be the random variable corresponding to flipping a coin, we may say X = 0 if the coin comes up heads, and X = 1 if the coin comes up tails. A finite number of outcomes makes the experiment discrete, whereas a continuum of outcomes makes it continuous.

Definition 1.1. *A Discrete Random Variable is a number that represents the outcome of some discrete random event.*

We can generalize about probability distributions, however, for the purposes of the question we analyze in this paper, discrete probability distributions are all we consider.

Definition 1.2. *A Discrete Probability Distribution is a function $f : X \to S$, for some $S \subset \mathbb{R}$ with finite or infinitely countable cardinality, which assigns a random variable, X, a number in S, with a known probability p_i, written $P(X = s_i) = p_i \; \forall \; s \in S$ and $p \in [0, 1]$, where $i \in \mathbb{N}$, that satisfies $\sum_{i=1}^{n} p_i = 1$.*

Taking the example of the coin flip, we have a random variable X, that is able to represent the situation. As we have set it up, $P(X = 0) = 0.5$ and $P(X = 1) = 0.5$. Note each probability is between 0 and 1, and the sum of the probabilities is 1. This sort of binary experiment is common and hence named.

Definition 1.3. *A Bernoulli Trial is an experiment in which there are only two outcomes, a "success" or a "failure." Success occurs with probability p, and failure occurs with probability (1-p), where $p \in [0, 1] \subset \mathbb{R}$.*

What if we wish to understand what happens when we flip a coin more than once? Suppose we want the probability of getting two heads out of three consecutive, independent coin tosses. This can happen three different ways. (i) Flipping heads twice, and then tails; (ii) Flipping heads, then tails, then heads; (iii) Flipping tails, and then heads twice. Each of these three outcomes has the same probability of occurring:

Example 1.1. *(i)* $P(X=0)*P(X=0)*P(X=1) = 0.5^2 * 0.5 = 0.125$,
(ii) $P(X=0)*P(X=1)*P(X=0) = 0.5*0.5*0.5 = 0.125$,
(iii) $P(X=1)*P(X=0)*P(X=0) = 0.5*0.5^2 = 0.125$.

Hence the total probability is P(exactly 2 heads out of 3 tosses) = 0.125 + 0.125 + 0.125 = 0.375.

Definition 1.4. *A Binomial Distribution assigns a random variable,* $Bin(n,p)$*, a number equal to the number of successes out of* $n \in \mathbb{N}$ *trials, where each trial is a Bernoulli Trial with a probability of success* p*.*

There is a strong intuitive understanding available to calculating the probability that a binomial random variable takes on any given value. For now:

(1) P(Bin(n,p) = k) = $\binom{n}{k}p^k(1-p)^{n-k}$, where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

When we ask about the probability of getting two heads out of three flips, we are really seeking P(Bin(3, 0.5) = 2). Using (1), that is simply $\binom{3}{2}p^2(1-p)^{3-2}$ = 0.375. What (1) says in symbols, is memorized easier by understanding where it comes from. Each sequence of Bernoulli trials that provides us the desired number of successes occurs with the same probability.

Example 1.2. *Success, Failure, Failure, Success, Success, Failure, Success =* $p(1-p)(1-p)pp(1-p)p = p^4(1-p)^3$ *= Four Successes followed by Three Failures*

In essence, you are multiplying the probability of getting k successes out of n trials, $p^{Successes}(1-p)^{Total\ Trials - Successes}$, times the number of sequences that have the probability of getting k successes, $\binom{n}{k}$.

1.2 Expected Value and Independent and Identically Distributed Random Variables

Any time we consider a probability distribution, we can discuss its expected value. Say we flip a coin six times and ask how many heads we can expect. Intuitively, as each trial has a 50% chance of being heads and a 50% chance of being tails, out of six trials, we can expect three of them to turn heads.

Definition 1.5. *The Expected Value of a discrete random variable, X, is* $E(X) = \sum_{i=1}^{n} P(X=x_i)*x_i$*, where n is the number of outcomes. In addition, the Expected Value of a transformation of X, g(X), where g sends X to a new random variable* $Y = g(X) = aX + b$ *for some* $a, b \in \mathbb{R}$*, is* $E(g(X)) = \sum_{i=1}^{n} P(X=x_i)*g(x_i)$*.*

Example 1.3. *If we let X be the random variable Bin(6, 0.5), assigning* x_i = *i heads, to obtain the expected number of heads for Bin(6, 0.5), we compute:* $\sum_{i=1}^{6} P(X=x_i)*x_i = P(X=x_1)*x_1 + P(X=x_2)*x_2 + ... + P(X=x_6)*x_6 = \binom{6}{1}0.5^1(1-0.5)^{6-1} + \binom{6}{2}0.5^2(1-0.5)^{6-2} + ... + \binom{6}{6}0.5^6(1-0.5)^{6-6} = 0.09375 + 0.46875 + ... + 0.09375 = 3$.

Theorem. *Let X be a binomial random variable, with parameters $n \in \mathbb{N}$ and $p \in [0,1] \subset \mathbb{R}$. The Expected Value of X is $E(X) = \sum_{i=1}^{n} P(X = x_i) * x_i = np$.*

Proof. We take the binomial theorem for real numbers to hold, without justification, which states that, for all $x, y \in \mathbb{R}$ and $n \in \mathbb{N}$, $(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^k$. Consider, then, $E(X) = \sum_{i=1}^{n} i \binom{n}{i} p^i (1-p)^{n-i} = \sum_{i=1}^{n} n \binom{n-1}{i-1} p^i (1-p)^{n-i} = np \sum_{i=1}^{n} \binom{n-1}{i-1} p^{i-1} (1-p)^{n-i} = np \sum_{i=1}^{n-1} \binom{n-1}{i} p^i (1-p)^{n-(i+1)}$. Let $k = i - 1$. Then $E(X) = np \sum_{i=1}^{n} \binom{n-1}{i-1} p^{i-1} (1-p)^{n-i} = np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-(k+1)}$ $= np(p + (1-p))^{n-1} = np(1)^{n-1} = np$ □

We can also consider the characteristics of a group of random variables. From here onward, we consider only Independent and Identically Distributed (IID) random rariables.

Definition 1.6 (IID Random Variables)**.** *Random variables who are mutually independent and come from the same distribution, with the same parameters, are said to be Independent and Identically Distributed.*

In the context of Binomial random variables, one can consider a new random variable $S_N = X_1 + X_2 + ... + X_N$, for $X_i = Binomial(n, p)$, with $n, N \in \mathbb{N}$ and $p \in [0,1]$, for all $\mathbb{N} \ni i \leq n$. Note that, in this case, we have $E(X_i) = E(X_j) \; \forall \; i, j \in [1, N] \subset \mathbb{N}$. This, in turns, allows us to denote the average of N random variables by $\bar{X} = \frac{S_N}{N}$.

Theorem. *The expected value of the average of N IID random variables, $\bar{X}$, is $E(\bar{X}) = E(X_i)$ for any $i \in [1, N] \subset \mathbb{N}$.*

Proof. We first verify that for all random variables, X, and for all $a \in \mathbb{R}$, we have $E(aX) = \sum_{i=1}^{n} P(X = x_i) * ax_i = a \sum_{i=1}^{n} P(X = x_i) * x_i = aE(X)$. Let $S_N = X_1 + X_2 + ... + X_N$ be the sum of N IID random variables for some $N \in \mathbb{N}$ and $\bar{X} = \frac{S_N}{N}$. Now consider $E(\bar{X}) = E(\frac{S_N}{N}) = \frac{1}{N} E(S_N) = \frac{1}{N} \sum_{i=1}^{N} E(X_i) = \frac{1}{N} N E(X_i) = E(X_i)$. This holds for any $i \in [1, N] \subset \mathbb{N}$, since having all X be IID necessitates $E(X_i) = E(X_j) \; \forall \; i, j \in [1, N]$. □

1.3 Sequences of Real Numbers, Monotonicity and Limits

Definition 1.7. *A sequence of real numbers, denoted $(x_n)_{n=1}^{\infty}$ or simply x_n, is a function $f : \mathbb{N} \to \mathbb{R}$.*

We have n=1,2,3,..., when, in reality, the domain for f can be replaced by $\{m \in \mathbb{Z} \; : \; m \geq n\}$, making $f : \mathbb{Z} \to \mathbb{R}$. The definition doesn't tell us much, so it may be useful to consider a few simple examples.

Example 1.4. $x_n = \frac{1}{n} = \{1, \frac{1}{2}, \frac{1}{3}, ...\}$; $y_n = n = \{1, 2, 3, ...\}$; $z_n = \frac{2n+1}{n} = \{3, \frac{5}{2}, \frac{7}{3}, ...\}$

For many sequences, as for the ones above, it's easy to tell whether a sequence is increasing or decreasing as n grows. For example, $x_n = \frac{1}{n}$ clearly grows smaller as n grows, and y_n and z_n are clearly increasing. From this we develop terminology to discuss such a phenomenon.

Definition 1.8. *We say a sequence of real numbers x_n is monotone decreasing if $x_n \geq x_{n+1}$ for all $n \in \mathbb{N}$, monotone increasing if $x_n \leq x_{n+1}$ for all n, and monotone if x_n is either monotone decreasing or monotone increasing.*

A particular case where x_n is the constant function, so that $x_n = c \in \mathbb{R}$ for some fixed c and all $n \in \mathbb{N}$, is monotone decreasing and monotone increasing. The constant sequence is monotone, which is enough to avoid issue.

Perhaps one notices that certain sequences tend towards some specific value as n increases. Consider $\frac{1}{n}$, which endlessly decreases in value towards some real number, which should be obvious. If you take n to be large, like 10,000, we get $x_{10,000} = \frac{1}{10000} = 0.0001$, a value very close to zero. Such intuition is formalized through the notion of limits.

Definition 1.9. *We say that a sequence x_n tends to a limit $L \in \mathbb{R}$, and we write $L = \lim\limits_{n\to\infty} x_n$, if for all $\mathbb{R} \ni \epsilon > 0$, there exists some $N \in \mathbb{N}$ such that whenever $n \geq N$, $|x_n - L| < \epsilon$.*

Theorem. *Let $x_n = \frac{1}{n}$. Then $L = \lim\limits_{n\to\infty} x_n = 0$.*

Proof. We take for granted that the Archimedian Property holds. The property states that for all $x \in \mathbb{R}$ there exists an $n \in \mathbb{N}$ such that $x < n$.
Fix $\epsilon > 0$. By the Archimedian Property, there exists $N \in \mathbb{N}$ satisfying $\frac{1}{\epsilon} < N$. Thus, $\frac{1}{N} < \epsilon$. Whenever $n \geq N$, we have that $\frac{1}{n} \leq \frac{1}{N} < \epsilon$, and so $\frac{1}{n} = |\frac{1}{n}| = |\frac{1}{n} - 0| < \epsilon$. Therefore $L = 0 = \lim\limits_{n\to\infty} x_n$. □

1.4 The Law of Large Numbers

From an informal standpoint, the Law of Large Numbers merely tells us that as we increase the number of trials toward infinity for some experiment, the probability that we get the obvious and expected result grows. For example, if we flip a coin six times and count the number of heads, the obvious and expected result would be three heads. The trial count is quite low, however, and the innate variance to such a procedure will likely introduce some error. Indeed, $E(Bin(6, 0.5)) = np = 6 * 0.5 = 3$, but $P(Bin(6, 0.5) = 3) = \binom{6}{3} 0.5^3 * (1 - 0.5)^{6-3} = 0.3125$, telling us that our expected result is not what what is likely to occur.

Theorem (The Law of Large Numbers)**.** *Let $X_1, X_2, ..., X_N$ be $N \in \mathbb{N}$ IID random variables, $\bar{X} = \frac{S_N}{N} = \frac{X_1+X_2+...+X_N}{N}$ be the average of these random variables, and $\mu = E(X_i)$ for any $i \in [1, N]$. Then for all $\mathbb{R} \ni \epsilon > 0$*
$P(|\bar{X} - \mu| < \epsilon) = 1$.

We omit the proof of the Law of Large Numbers, and leave in its stead the discussion above of the intuition surrounding it.

2 Introduction

Our question revolves entirely around considering the probability that the random variable Bin(n, p) is greater than or equal to some portion of n. For example, we can calculate the probability that at least 80% of 10 coin flips are heads, or, equivalently, the probability of flipping a coin 10 times resulting in 8, 9, or 10 heads.

Example 2.1. *P(8, 9, or 10 heads out of 10 coin flips) = P(Bin(10, 0.5) $\geq$ 8) = P(Bin(10, 0.5) = 8) + P(Bin(10, 0.5) = 9) + P(Bin(10, 0.5) = 8)* $= \binom{10}{8}0.5^8(1-0.5)^{10-8} + \binom{10}{9}0.5^9(1-0.5)^{10-9} + \binom{10}{10}0.5^10(1-0.5)^{10-10} =$ *0.0439453125 + 0.009765625 + 0.0009765625 = 0.0546875*

Definition 2.1. *Let p,q* $\in [0,1] \subset \mathbb{R}$ *and* $n \in \mathbb{N} = \{1, 2, 3, ...\}$*. We define the function* $A : \mathbb{R}^2 \times \mathbb{N} \to \mathbb{R}$ *by* $A(p, q, n) = P(Bin(n, p) \geq qn)$*.*

This definition of A provides us a convenient way to analyze the probability of a binomial random variable being at least as large as some portion of n. For instance, if we take q = $\frac{3}{4}$, then we seek to understand the probability that, Bin(n,p) $\geq \frac{3}{4}n$. We are now in a position to reveal our initial conjecture.

Conjecture 2.1. *Let* $q = \frac{3}{4}$*, then A(p,q,4l) is a monotone sequence.*

The hope in establishing this simple conjecture was to find a case which would provide motivation for believing a stronger, more general version. Conjecture 2.1 was quickly disproven, however.

Example 2.2. *Let* $q = \frac{3}{4}$*,* $p = 0.8$*, and consider A for* $l = 2, 3, 4$*.* $A(0.8, \frac{3}{4}, 8) = P(Bin(8, 0.8) \geq \frac{3}{4}8) = P(Bin(8, 0.8) \geq 6) = 0.79691776$*. Likewise,* $A(0.8, \frac{3}{4}, 12) = 0.7945689497600001$*, and* $A(0.8, \frac{3}{4}, 16) = 0.7982454417653762$*. Hence, we have* $A(0.8, \frac{3}{4}, 16) > A(0.8, \frac{3}{4}, 8) > A(0.8, \frac{3}{4}, 12)$*, violating* $A(0.8, \frac{3}{4}, 4l)$*'s monotonicity.*

3 Modern Developments

After the initial conjecture was falsified, efforts were to be spent elsewhere, however a more restrictive, and less precise, conjecture was proposed.

Conjecture 3.1. *Let* $q = \frac{\alpha}{\beta}$ *where* $\alpha, \beta \in \mathbb{N}$*, and let* $k \in \mathbb{N}$*. Then there exists an* $\epsilon \in \mathbb{R}$*, with* $\epsilon > 0$*, such that whenever* $|p - q| \geq \epsilon$*,* $A(p, q = \frac{\alpha}{\beta}, k\beta)$ *is monotone.*

If we henceforth assume that $p > q$, the Law of Large Numbers assures us that as $n \to \infty$, $A(p, q, n) \to 1$. Then it may be fruitful to find for which n we have $A(p, q, n+1) < A(p, q, n)$. From this question we discovered a pattern, but first, the setup.

Definition 3.1. *Let $x \in \mathbb{R}$. Then we define the floor of x to be the largest integer less than or equal to x. Let $z \in \mathbb{Z}$ be such an integer, so that $x \leq z$. Then we write $\lfloor x \rfloor = z$. The ceiling of a real number x is the smallest integer greater than or equal to x. Let $a \in \mathbb{Z}$ be such an integer, so that $a \geq x$. Then we write $\lceil x \rceil = a$.*

The floor of a real number is useful to us as it allows us to find the places where $A(p, q, n+1) < A(p, q, n)$, or $P(Bin(n+1, p) \geq q(n+1)) < P(Bin(n, p) \geq qn)$. Intuitively this would only be true whenever $\lceil q(n+1) \rceil > \lceil qn \rceil$. To see why, consider, for example, tossing a coin 16 times versus 15 times and counting the heads. If we ask the question: Am I more likely to get at least 8 heads out of 15 coin flips or at least 8 heads out of 16 coin flips?

Example 3.1. *Let $p = 0.5$, $q = 0.5$, $n = 15$. We are really comparing $P(Bin(15, 0.5) \geq 8)$ versus $P(Bin(16, 0.5) \geq 8)$, since $\lceil q(n+1) \rceil = \lceil 0.5(16) \rceil = 8 = \lceil 0.5(15) \rceil = \lceil qn \rceil$.*

$$A(0.5, 0.5, 15) = 0.5$$

$$A(0.5, 0.5, 16) = 0.5981903076171875$$

The only time that we should see $A(p, q, n+1) < A(p, q, n)$, then, is when $\lceil q(n+1) \rceil > \lceil qn \rceil$. Do note, however, that this doesn't mean that $P(Bin(n+1, p) = k)$ has any business being greater than or equal to $P(Bin(n, p) = k)$ in many cases.

Example 3.2. *Let $p = 0.5$ and $n = 15$. Observe that $P(Bin(16, 0.5) = 7) = 0.17456054688 > P(Bin(17, 0.5) = 7) = 0.14837646484$.*

Example 3.2 is to help distinguish the notion of a binomial random variable equalling some number versus a binomial random variable being ***greater than or equal to some portion of n***.

Returning to the question of when $A(p, q, n+1) < A(p, q, n)$, we define a number that tells us precisely this.

Definition 3.2. *Let $p > q = \frac{\alpha}{\beta}$. We define the increment points as $n_k = \left\lfloor \frac{k}{q} \right\rfloor = \left\lfloor \frac{\beta q}{\alpha} \right\rfloor$.*

It as at these points, n_k, that $\lceil q(n+1) \rceil > \lceil qn \rceil$, and so we have good reason to believe that whenever $\frac{k}{q} \neq \left\lfloor \frac{k}{q} \right\rfloor$ then $A(p, q, n+1) < A(p, q, n)$. Defining increment points leads to some neat patterns.

Example 3.3. *Let $q = 3/17$ so that $n_k = \left\lfloor \frac{\beta k}{\alpha} \right\rfloor = \left\lfloor \frac{17k}{3} \right\rfloor$.*

$$n_1 = \lfloor 17/3 \rfloor = 5, n_2 = \lfloor 34/3 \rfloor = 11, n_3 = \lfloor 51/3 \rfloor = 17$$

With a longer sequence being

$$5, 11, 17, 22, 28, 34, 39, 45, 51, 56, 62, 68, 73, 79, 85, 90, 96, 102, 107, 113, ...$$

Inserting a 0 at the beginning of the sequence and taking the difference of consecutive terms

$$5, 6, 6, 5, 6, 6, 5, 6, 6, 5, 6, 6, \ldots$$

Remark. *(i) We have that* $n_\alpha = \left\lfloor \frac{\beta\alpha}{\alpha} \right\rfloor = \beta, n_{2\alpha} = \left\lfloor \frac{\beta 2\alpha}{\alpha} \right\rfloor = 2\beta$*, and, in general,* $n_{k\alpha} = k\beta$*.*
(ii) Write, for $x \in \mathbb{R}$*,* $\langle x \rangle = x - \lfloor x \rfloor$*. Then*

$$n_k = n_{\alpha \lfloor \frac{k}{\alpha} \rfloor} + n_{\langle \frac{k}{\alpha} \rangle} = \left\lfloor \frac{k}{\alpha} \right\rfloor \beta + n_{\langle \frac{k}{\alpha} \rangle} \tag{1}$$

Example 3.4. *Let* $q = 3/17$*, and it is known that* $n_2 = 11$*. With this, compute* n_{20}*.*

$$n_{20} = n_{3 \lfloor \frac{20}{3} \rfloor} + \ n_{\langle \frac{20}{3} \rangle} = n_{18} + \ n_{\frac{20}{3} - \lfloor \frac{20}{3} \rfloor} = \left\lfloor \frac{20}{3} \right\rfloor 17 + n_2 = 102 + 11 = 113$$

Indeed, $n_{20} = \left\lfloor \frac{20*17}{3} \right\rfloor = \left\lfloor \frac{340}{3} \right\rfloor = 113$*.*

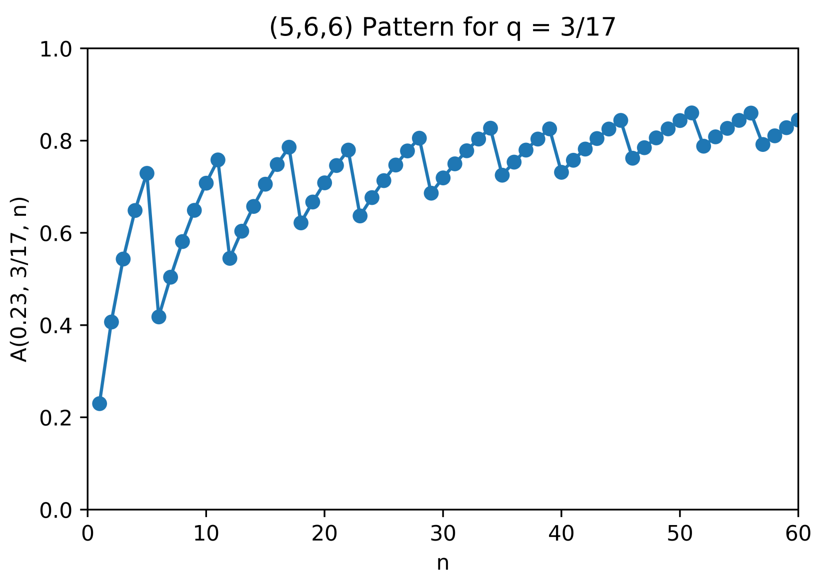
If you take (1) without justification, then we believe it implies the repeating pattern of the sort observed in Example 3.3. In addition, we are now ready to reveal a new theorem.

Theorem. *Let* $p > q = \frac{\alpha}{\beta}$ *for* $\alpha, \beta \in \mathbb{N}$*, and define* $n_k = \left\lfloor \frac{k}{q} \right\rfloor$*. Then*

$$A(p, q, n+1) < A(p, q, n) \implies n = n_k \ \exists \ k \in \mathbb{N}$$

.

An illustrative example is shown in the figure below:



Proof. Proof pending... (in notebook currently) $\square$