

# Analyzing Outliers and Significance: Insights into Dunbar's Number and Genetic Conditions in Human Evolution

Mina Lu

Received January 07, 2024

Accepted April 04, 2024

Electronic access April 15, 2024

This article introduces the Square Root Law of Outliers and provides its derivation, highlighting its significance and the context of its application. It explores the use of the Inverse of the Herfindahl-Hirschman Index (IHHI) in analyzing data distributed across bins in a normal distribution, and establishes a natural level of significance for the Student's t-test in the limit of large sample sizes. This approach leads to the identification of a critical value  $t_{\text{critical}} = \sqrt{\pi} = 1.77$ , and a two-sided natural level of significance  $\alpha = 0.076$ . By combining the Square Root Law of Outliers with this natural level of significance in an analysis of how humans make friends, and by applying them to a "likability" measure, the article successfully formulates an analytical expression for Dunbar's Number. This offers a mathematical explanation for the cognitive limits of human social networks. As a concrete application of the Square Root Law of Outliers, and by interpreting Dunbar's Number as the typical size of a hunter-gatherer community, the study analyzes the prevalence rates of genetic conditions such as color-blindness, autism, and bipolar disorder. It paves the way for a novel hypothesis: that these genetic conditions were integrated into the human genetic pool approximately 10,000 years ago in early agricultural societies, serving as evolutionary advantages in warfare. The article also discusses some other potential areas of application for the Square Root Law of Outliers.

## Introduction

The "Square Root Law" presents itself in many situations. The most famous example is the so-called "Birthday Problem<sup>1</sup>." This problem consists in estimating the necessary number of people for at least two of them to share a birthday, with a probability greater than 0.5. The correct answer is 23 people, far less than the number of days in a year, but closer to  $\sqrt{365} \approx 19.1$ . Another well-known situation where a square root law appears is in the Central Limit Theorem<sup>2</sup>, which states that given  $N$  i.i.d. (independent and identically distributed) random numbers each with mean  $\mu$  and standard deviation  $\sigma$ , the distribution for the average value of these  $N$  random variables will approach a normal distribution with the same mean  $\mu$  but with a much narrower standard deviation  $\sigma/\sqrt{N}$ . A square root law also shows up in Brownian motion<sup>3</sup>, where the displacement is expected to increase as the square root of time. Square root laws have also been applied empirically to various contexts, such as in (a) Inventory management<sup>4</sup>, where it states that the number of warehouses only needs to increase as the square root of size of inventory, in (b) in analysis of productivity, where it is known as Price's Square Root Law<sup>5</sup>, stating that 50% of the results will be achieved by the square root of the size of the total group of workers, or in (c) pharmaceutical inspection, where it is known as the rule of "Square Root of  $N$  Plus One"<sup>6</sup>  $\sqrt{N+1}$ , and it dictates how many boxes to sample and inspect.

There appears to be no published results on the Square Root

Law as pertaining to the size of outliers, and this article represents an effort to provide a context and justification for its applicability in this new domain.

Establishing a Square Root Law of Outliers aids in situations where attention comes at a premium. For instance, in statistical analysis, continuous variables are often binned into classes for ease of discussion. An example is the use of tax brackets. Also, in the plotting of histograms, often the last category is an exceptional category like "\$100 or above." The use of the Square Root Law of Outliers helps to find an appropriate cutoff point for defining outliers.

This law also contributes to the justification of the distribution of resources in policymaking. For instance, many diseases warrant research studies. However, governments and human societies have limited resources. Therefore, it becomes a matter of practical necessity to fund only those research studies on diseases that impact a significant portion of the population. Similarly, a political representative (e.g., a city council member or a congressperson) cannot possibly address all the concerns of their individual constituents and must focus only on those matters that are of general concern. The Square Root Law also helps to establish the selection criterion and separate the exceptional members from regular members. For instance, there are almost 200 countries in the world, which gives a square root value of  $\sqrt{200} \approx 14.1$ . Therefore, it is not surprising that a forum like the G8 (The Group of Eight) or G7 (The Group of Seven) has been criticized as being too narrow, and a more inclusive forum

like the G20 (The Group of Twenty) has emerged to allow for a more comprehensive dialogue on issues related to the global economy.

In the context of evolution, genetic variations may either bring about novel subspecies/species or cause genetic anomalies. The Square Root Law of Outliers helps to separate normal variations from anomalies and may assist in the choice of action between accommodation and treatment.

## The Square Root Law of Outliers

This new square root law applies to competition between two teams: team I and team II. First, assume team I is composed of i.i.d. team members, each capable of making a contribution that follows a distribution with mean  $\mu_A$  and standard deviation  $\sigma_A$ . By the Central Limit Theorem, the average member contribution in team I would approach a normal distribution with mean  $X_I = \mu_A$ , but with a standard deviation approaching  $\sigma_I = \frac{\sigma_A}{\sqrt{N}}$ , where  $N$  is the size of team I.

Now, consider a second team II with the same sample size  $N$ , but with two different types of team members:

- $(N - n)$  i.i.d. team members just like those in team I, and
- $n$  i.i.d. team members of a different kind that make a contribution following a distribution with a mean  $\mu_B > \mu_A$  and a standard deviation  $\sigma_B$ .

Now, the two teams will compete. The question is: which team will win? Obviously, team II has an advantage because  $\mu_B > \mu_A$ . But in real life, a better team may still lose to a worse team, due to random factors. The question is then: at what size of  $n$  would it be most clear that team II will win in a game against team I? This question mimics the Student's  $t$ -test<sup>7</sup> for comparing two means. Essentially, the new mean  $X_{II}$  should differ from the old one  $X_I$  by an amount larger than the standard error coming from statistical fluctuations. In this test, one defines a  $t$ -statistic to be:

$$t = \frac{X_{II} - X_I}{\sigma}$$

$$\sigma = \sqrt{\sigma_I^2 + \sigma_{II}^2}$$

where a sample of size  $N$  for each team has been assumed, as represented by their respective  $N$  team members. In that case,

the following formulas result:

$$X_I = \mu_A$$

$$\sigma_I = \frac{\sigma_A}{\sqrt{N}}$$

$$X_{II} = \frac{(N - n)\mu_A + n\mu_B}{N} = \mu_A + \frac{n}{N}(\mu_B - \mu_A)$$

$$\sigma_{II} = \frac{1}{N} \sqrt{(N - n)\sigma_A^2 + n\sigma_B^2} = \frac{\sigma_A}{\sqrt{N}} \sqrt{1 + \frac{n}{N} \left( \frac{\sigma_B^2}{\sigma_A^2} - 1 \right)}$$

From there, it can be inferred that

$$\sigma = \sqrt{\sigma_I^2 + \sigma_{II}^2} = \frac{\sigma_A}{\sqrt{N}} \sqrt{2 + \frac{n}{N} \left( \frac{\sigma_B^2}{\sigma_A^2} - 1 \right)}$$

$$t = \frac{X_{II} - X_I}{\sigma} = \frac{\frac{n}{N}(\mu_B - \mu_A)}{\frac{\sigma_A}{\sqrt{N}} \sqrt{2 + \frac{n}{N} \left( \frac{\sigma_B^2}{\sigma_A^2} - 1 \right)}} = \frac{n}{\sqrt{N}} \frac{(\mu_B - \mu_A)}{\sigma_A \sqrt{2 + \frac{n}{N} \left( \frac{\sigma_B^2}{\sigma_A^2} - 1 \right)}}$$

In the limit  $N \gg n \gg 1$ , the  $t$  statistic becomes

$$t = \frac{n}{\sqrt{N}} \frac{(\mu_B - \mu_A)}{\sigma_A \sqrt{2}}$$

In the limit of a large number of degrees of freedom  $N - 1$  in the  $t$ -test, the Student's  $t$ -distribution approaches a normal distribution  $N(0, 1)$ . Thus, one simply looks up the inverse cumulative distribution function of  $N(0, 1)$  to obtain a desired critical value. (In the SciPy Python package, this is given by the `scipy.special.ndtri()` method). For instance, at  $\alpha = 0.05$  significance level,  $t_{\text{critical}} = 1.64$ . Using the critical value for a given significance level, one can then infer from the above formula the value of  $n_{\text{critical}}$  necessary to more-or-less guarantee team II winning a match against team I. That is:

$$n_{\text{critical}} = \sqrt{N} \frac{(t_{\text{critical}} \sigma_A \sqrt{2})}{(\mu_B - \mu_A)}$$

As seen in the next section, there is a natural choice of for the critical value given by  $t_{\text{critical}} = \sqrt{\pi} = 1.77$ , corresponding to a singled-sided significance level of  $\alpha = \frac{1}{2} \text{erfc} \left( \sqrt{\frac{\pi}{2}} \right) = 0.038$ . Choosing this value of  $t_{\text{critical}}$ ,

$$n_{\text{critical}} = \sqrt{N} \frac{(\sigma_A \sqrt{2\pi})}{(\mu_B - \mu_A)}$$

Assuming the factor inside the parenthesis to be around 1, a general rule is created:

$$n_{\text{critical}} \sim \sqrt{N}$$

That is, a subpopulation of high-performing outliers will only become statistically significant if their size is larger than the

square root of the total population. Incidentally, by taking the natural choice of  $t_{\text{critical}} = \sqrt{\pi} = 1.77$ , if the multiplicative factor in the square root law were to be 1, that would mean:

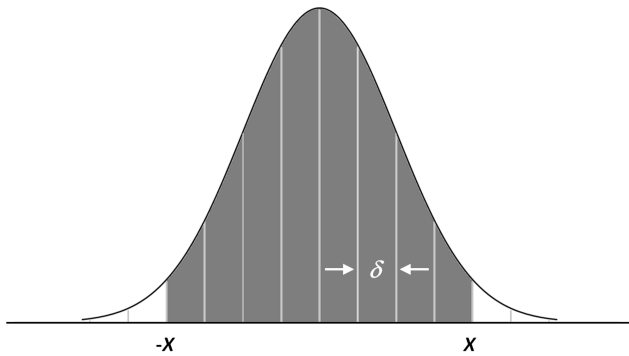
$$\frac{(\mu_B - \mu_A)}{\sigma_A} = \sqrt{2\pi} = 2.51$$

That is, the new mean  $\mu_B$  would be about two and half times larger than the standard deviation, away from the mean of the distribution of contribution of a member in team I. This is a reasonable requirement for the new type of team member to excel over the existing type. Of course, if this is not the case, the multiplicative factor in the square root law will need to be adjusted accordingly.

Obviously, this analysis carries through for the case of underperforming outliers as well. In general, in a population of  $N$  members, due to the square root law and under fairly general conditions, usually a subgroup's differential contribution will stand out only if the size of the subgroup is larger than  $\sqrt{N}$ , save for a multiplicative factor that is typically of the order of 1.

### Natural Level of Significance $\alpha = 0.076$

Whenever one performs a Student's  $t$ -test, one must always specify a level of significance  $\alpha$ . Popular choices of level of significance are 0.10, 0.05, and 0.01. One may wonder whether there is a more natural way of automatically selecting a level of significance that is suitable to most typical situations. By applying the *Inverse Herfindahl-Hirschman Index* (IHHI)<sup>8</sup> to a binned normal distribution, a natural value for a level of significance can be effectively attained.



**Fig. 1** Level of significance

The continuous distribution of  $N$  samples that follows a normal distribution  $N(0, \sigma^2)$  is given by

$$f(x) = \frac{N}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

Upon binning, the population size in each bin (using a left Riemann sum) is given by:

$$n_i = \frac{N\delta}{\sigma\sqrt{2\pi}} e^{-\frac{x_i^2}{2\sigma^2}} = \frac{N\delta}{\sigma\sqrt{2\pi}} e^{-\frac{i^2\delta^2}{2\sigma^2}}; \quad i = -\infty, \dots, -2, -1, 0, 1, 2, \dots, +\infty$$

The “effective number of bins” can be obtained by calculating the Inverse Herfindahl-Hirschman Index (IHHI), which is also commonly used in the analysis of market concentration and monopoly situations<sup>9</sup>. It has also been applied to the effective number of political parties in parliaments<sup>8</sup>.

$$B = \text{effective number of bins} = \frac{N^2}{\sum_i n_i^2}$$

In the continuous limit  $\delta \rightarrow 0$ ,

$$B = \frac{2\pi\sigma^2}{\delta^2} \frac{1}{\sum_i e^{-\frac{i^2\delta^2}{\sigma^2}}} \approx \frac{2\sigma\sqrt{\pi}}{\delta} \frac{1}{\left(\frac{1}{\sigma\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{\sigma^2}} dx\right)} = \frac{2\sigma\sqrt{\pi}}{\delta}$$

When assigning these  $B$  bins to the central area of the normal distribution, the total horizontal span of the shaded area would be  $2X = B\delta = 2\sigma\sqrt{\pi}$ . This means

$$X = \sigma\sqrt{\pi} = 1.77\sigma$$

The effective fraction of the total population (the shaded area) would be:

$$P_{\text{effective}} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\sigma\sqrt{\pi}}^{\sigma\sqrt{\pi}} e^{-\frac{x^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{\pi}}^{\sqrt{\pi}} e^{-\frac{x^2}{2}} dx = \text{erf}(\sqrt{\pi/2}) = 0.924 = 92.4\%$$

where  $\text{erf}(t)$  is the error function. The two-sided level of significance can be expressed in terms of the complement of the error function  $\text{erfc}(t)$ :

$$\alpha_{\text{two-sided}} = 1 - P_{\text{effective}} = 1 - \text{erf}(\sqrt{\pi/2}) = \text{erfc}(\sqrt{\pi/2}) = 0.076$$

whereas the single-sided level of significance would be given by

$$\alpha_{\text{single-sided}} = \frac{1}{2} \text{erfc}(\sqrt{\pi/2}) = 0.038$$

The corresponding critical value for a Student's  $t$ -statistic in the limit of large  $N$  is given by:

$$t_{\text{critical}} = \sqrt{\pi} = 1.77$$

One may wonder why bother choosing a natural level of significance or a critical value of the  $t$ -statistic. This question is similar to asking why one should choose Euler's number  $e = 2.71828\dots$  as the base of the natural logarithmic function, when any other base would do. This article aims to prove that the special analytic properties of these natural choices are what ultimately lend support to their usage.

---

## An Analytical Expression for Dunbar's Number

In the ambit of social sciences, Dunbar's Number<sup>10</sup> has widely been used as a suggested cognitive limit to the number of people with whom one can maintain stable social relationships. It was first introduced by the British anthropologist Robin Dunbar, and it proposes that humans can maintain 150 stable relationships. Dunbar has also analyzed the typical size of hunter-gatherer societies. Dunbar's number has been proposed to lie between 100 and 250, with a commonly used value of 150<sup>11</sup>.

Although Robin Dunbar reached his number by studying relations between primate brain sizes and their average social group size, one may wonder whether there is a more fundamental mathematical mechanism behind this number. This section represents such an attempt.

The creation of a measure of "likability" of a person's friends serves as a foundational step in this approach. A friend may score high if they are a candidate for a romantic relationship. Another friend may score high if they are a good fellow co-worker. However, romantic relationships are typically discouraged in a workplace environment. That is, a person that is likable in one dimension may not be likable in another dimension. No one person can score high in all dimensions simultaneously. There are a multitude of dimensions to measure a person's likability. A Principal Component Analysis (PCA)<sup>12</sup> can be applied to synthesize all these dimensions down to one single dimension. Once typical variable centering and z-scaling<sup>13</sup> are performed, the distribution of people that a person regularly interacts with is expected to more-or-less follow a normal distribution, along the direction of the first principal component. It is important to point out that people that a person may like a lot may not reciprocate, since they themselves may prefer to hang around with other friends that they like better.

By applying the "Inverse Herfindahl-Hirschman Index" analysis of the previous section, it can be inferred that out of the  $N$  friends that a person regularly interacts with, two groups may emerge:

- Stable friends: this group constitutes  $\operatorname{erf}\left(\sqrt{\frac{\pi}{2}}\right) = 0.924 = 92.4\%$  of friends, with whom a person has stronger attachments to, and these friends are thus within the "effective bins" in the normal distribution. They would fall under the "exploitation" category.
- Outlier friends: this group constitutes  $\operatorname{erfc}\left(\sqrt{\frac{\pi}{2}}\right) = 0.076 = 7.6\%$  of friends, with whom a person has weaker attachments to, and these friends are thus outside the "effective bins" of the normal distribution. These are the outlier friends. They would fall under the "exploration" category.

If the square root law from the first section of this article is further applied to this dichotomy of friends, the size of outliers can be estimated to be  $\sqrt{N}$ , which is the number of friends that

a person doesn't particularly like (left-tail region of the normal distribution), or that they are not particularly liked by (right-tail region of the normal distribution). That is, the fraction of these outlier friends will be:

$$\operatorname{erfc}\left(\sqrt{\frac{\pi}{2}}\right) = \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}}$$

which means that the number of friends that simultaneously satisfy the IHHI criterion and the square root law of outliers is given by

$$N = \frac{1}{\left(\operatorname{erfc}^2\left(\sqrt{\frac{\pi}{2}}\right)\right)} = 171.7$$

In particular, the size of stable friends (the "exploitation" group), which shall be identified as the Dunbar's Number  $D$ , is given by

$$\begin{aligned} D &= N - \sqrt{N} = N \left(1 - \frac{\sqrt{N}}{N}\right) = \frac{(1 - \operatorname{erfc}\left(\sqrt{\frac{\pi}{2}}\right))}{\left(\operatorname{erfc}^2\left(\sqrt{\frac{\pi}{2}}\right)\right)} \\ &= \frac{\operatorname{erf}\left(\sqrt{\frac{\pi}{2}}\right)}{\operatorname{erfc}^2\left(\sqrt{\frac{\pi}{2}}\right)} = 158.6 \end{aligned}$$

The numerical value obtained here is fairly close to the value 150 that is commonly used for Dunbar's number.

How should one interpret the origin of this number? This present paper obviously has not looked into the size of brains of primates—it has only relied on a mathematical analysis. Yet, this analysis finds a number that is in the general vicinity of the commonly used value. Dunbar's number may be a natural consequence, if the human brain manages relationships in two layers: first at an individual level, and then at a group level. This is akin to how a country like the U.S.A. is organized: it has a federal government, but the country is also divided into states. This two-level split may ultimately be the responsible for the occurrence of Dunbar's Number. That is, if a person correlates and associates their friends into groups so they can better manage their relationships inside their brain, then Dunbar's Number could arise as a natural consequence. As to why other primates do not reach the level of Dunbar's Number in the number of their relationships, a possible explanation is that they have not fully developed this two-layer split inside their brains.

## An Application

The Square Root Law of Outliers is applicable to any situation of competition. For instance, in the context of evolution; due to recombination, natural mutations and epigenetic effects; genetic expressions of individuals will differ. Unusual genetic conditions may be interpreted as Mother Nature doing exploration instead of exploitation. This begs the question: what types of genetic conditions should be considered normal variations, and

---

what types of genetic conditions should be considered outliers? The Square Root Law of Outliers provides one simple rule: if a genetic subgroup's prevalence rate is above  $\frac{1}{\sqrt{N}}$ , then the condition is considered to be a part of normal variation, whereas if a genetic subgroup's prevalence rate is below  $\frac{1}{\sqrt{N}}$ , then it can be regarded as an outlier. Here,  $N$  represents a collective group of a population that competes with other collective groups. Within each collective group, assume that the group members may reproduce randomly with other members. In human society, the unit of competition typically is at the level of countries, because within the same country, people can move freely and form families more easily. In antiquity, the unit of competition could be human clans or tribes. Additionally, wars often happen between different clans, which provide a force of natural selection. The typical size of a country today is around 10 million people. That would mean that any genetic subgroup with a prevalence rate above  $\frac{1}{\sqrt{10,000,000}} \approx 0.03\%$  should be considered as a significant contributor. That is, such a subgroup probably carries some evolutionary reason for its existence.

Autism has a prevalence rate of 1 in 36 children<sup>14</sup>, or about 3%. This prevalence rate is two orders of magnitude higher than the threshold value of 0.03%. Therefore, there are probably some good evolutionary reasons why Mother Nature has included this subgroup in humans. The same is true with color blindness, which makes up about 2% to 4% of the general population (4% to 8% in boys)<sup>15</sup>. The same is also true with bipolar disorder (lifetime prevalence rate 4.4% in the U.S. adults)<sup>16</sup>. On the other hand, congenital hypothyroidism has an incidence rate of 1:3000 to 1:4000, or below 0.03%, making it suitable to be considered an anomaly<sup>17</sup>. A condition such as Phelan-McDermid Syndrome has an occurrence rate of 2.5–10 per million births<sup>18</sup>, or around 0.0005%, so it should be considered an anomaly as well. This separation between normal variations and anomalies can guide the creation of an approach to handling these different types of population subgroups. In fact, in the case of color blindness, modern society has largely accepted and accommodated color-blind citizens. In the case of autism, instead of “treatments,” perhaps others should try to explore the reasons behind Mother Nature’s design, and provide alternative development paths for children with this genetic condition.

The prevalence rate of color-blindness, autism and bipolar disorder all fall around the general vicinity of 3%–4%. One may wonder whether this is beyond just a coincidence. If Dunbar’s Number  $\sim 150$  indeed is the size of a typical hunter-gatherer community, if the square root law is applied to the inverse of the prevalence rate, then the critical population size for the three genetic conditions (color-blindness, autism and bipolar disorder) would be around  $(\frac{1}{0.04})^2 = 625$ , or about the size of early agricultural villages.

There are anecdotal accounts on the ability of color-blind soldiers being able to see through enemy camouflage. A more

formal study has confirmed the ability of color-blind people to see through camouflage<sup>19</sup>. Camouflage is particularly important after the invention of long-range projectile weapons, such as bows and arrows. Given that color-blindness is predominantly a male-bound genetic condition, it seems like the linkage of color-blindness to a combat advantage is at least plausible, since in antiquity, combat roles were mostly male-bound. Similarly, autism occurs more often in boys than girls (about 3:1 according to one study<sup>20</sup>), although less clear-cut as the case of color blindness. Given the gender bias of color blindness and its relationship to a warfare advantage, there is a possibility that autism also provided an edge in ancient battles. On the other hand, bipolar disorder shows no gender preference. Putting all these three genetic conditions together, it seems to suggest that in antiquity, color-blind people had a direct role in combat situations, whereas the autistic group more likely played a supporting role, and that the bipolar group mostly helped combat efforts away from battlefields. Today, many autistic people work in the fields of science, technology, accounting, etc., whereas bipolar disorder is well known within the worlds of artists and performers, whose skills may come in handy in rallying the population behind supporting war efforts. So, an idea can be formed about the roles of these groups of people during warfare. Obviously, not all color-blind people are warriors, nor are all autistic people engineers, nor are all bipolar-disorder people artists or performers. These subgroups instead come as statistical distributions, and it is their overall distributions as subgroups that gave an edge to their clans. Given the prevalence rates of these three groups, perhaps these genetic conditions entered the human race to help with the survival of individual tribes. Tribes without these genetic traits may have had disadvantages in combat situations, and if natural selection is to hold, those tribes may have perished in history. That is, modern human societies may all be carriers of these three genetic conditions. Furthermore, after the prevalence rates of these three conditions cross over their critical values (due to warfare and elimination of tribes not possessing these genetic traits), all surviving tribes become equally endowed with these subgroups of people and these genetic conditions cease to be advantages. That is, all surviving groups have the exact same advantage. Therefore, it is possible to have a “decoupling” situation, whereby the prevalence rates remain “frozen” even as societal sizes increase. Although this conjecture is speculative, it holds importance and deserves to be described. If this conjecture holds true, then it suggests that color blindness, autism and bipolar disorder entered the human race at around the time of transition from hunter-gatherer society to the Neolithic agricultural society, or roughly 10,000 years ago. Though violence has always existed in human societies, many researchers believe that organized warfare is an invention of the Neolithic period<sup>21</sup>.

This conjecture may be testable in the future, if more DNA samples from early humans can be recovered. Even without

---

historical DNA evidences, studies on modern humans may also shed light on the plausibility of this conjecture. In the case of color-blindness, some studies have been conducted a few decades ago on surviving hunter-gatherer societies in the world<sup>22</sup>, and researchers have generally found that color blindness is rare in these populations. Similarly, there is also some evidence that bipolar disorder may be rare among hunter-gatherers<sup>23</sup>. All these findings lend support to this conjecture. It is perhaps not as easy today to conduct similar research regarding autism, due to the dwindling number of hunter-gatherer societies worldwide. Nonetheless, this remains an interesting topic to explore, as many researchers are looking at these genetic conditions from the perspective of evolution.

## Conclusion

In conclusion, this paper has introduced a Square Root Law applicable to the estimation of size of outliers, and stated the context of its derivation. By combining this law with the application of the Inverse Herfindahl-Hirschman Index to the binned normal distribution, an analytic expression for Dunbar's Number is created. The Square Root Law of Outliers is proposed to be a simple rule of thumb to distinguish normal variations from anomalies for subgroups of genetic conditions. By applying this analysis to genetic conditions such as color blindness, autism and bipolar disorder, this paper proposes that these conditions may have entered the human race about 10,000 years ago to provide advantages in warfare. The application of the Square Root Law of Outliers extends beyond human genetics. It can be applied to general situations where limited attention or resources pose a constraint, such as prioritizing disease research studies or addressing political concerns from constituents. It can also be used to find an appropriate cutoff value for the definition of exceptional categories.

## References

- 1 Wikipedia, *Birthday problem*, [https://en.wikipedia.org/wiki/Birthday\\_problem](https://en.wikipedia.org/wiki/Birthday_problem).
- 2 Wikipedia, *Central limit theorem*, [https://en.wikipedia.org/wiki/Central\\_limit\\_theorem](https://en.wikipedia.org/wiki/Central_limit_theorem).
- 3 Wikipedia, *Brownian motion*, [https://en.wikipedia.org/wiki/Brownian\\_motion](https://en.wikipedia.org/wiki/Brownian_motion).
- 4 D. Maister, *International Journal of Physical Distribution*, **6**, 124–134.
- 5 D. J. Solla Price, *Little Science, Big Science – and Beyond*.
- 6 H. Saranadasa, *Pharmaceutical Technology*, **27**, 50–62.
- 7 Wikipedia, *Student's t-test*, <https://en.wikipedia.org/wiki/Student%27s-t-test>.
- 8 Wikipedia, *Effective number of parties*, [https://en.wikipedia.org/wiki/Effective\\_number\\_of\\_parties](https://en.wikipedia.org/wiki/Effective_number_of_parties).
- 9 Wikipedia, *Herfindahl–Hirschman index*, [https://en.wikipedia.org/wiki/Herfindahl%E2%80%9393Hirschman\\_index](https://en.wikipedia.org/wiki/Herfindahl%E2%80%9393Hirschman_index).
- 10 R. Dunbar, *Journal of Human Evolution*, **22**, 469–493.
- 11 D. Wikipedia.
- 12 I. Jolliffe and J. Cadima, *Phil Trans. R. Soc.*, **374**, year.
- 13 Wikipedia, *Standard score*, [https://en.wikipedia.org/wiki/Standard\\_score](https://en.wikipedia.org/wiki/Standard_score).
- 14 M. J. Maenner, *United States*, **72**, 1–14.
- 15 J. Birch, *Journal of the Optical Society of America A*, **29**, 313–320.
- 16 K. R. Merikangas, H. S. Akiskal and J. Angst, *Arch Gen Psychiatry*, **64**, 543–552.
- 17 L. Liu, W. He and J. Zhu, *Eur J Pediatr*, **182**, 2957–2965.
- 18 F. Cammarata-Scalisi, *Genes (Basel)*, **13**, 504.
- 19 M. Morgan, A. Adam and J. D. Mollon, *Proc. R. Soc. Lond. B*, **248**, 291–295.
- 20 R. Loomes, L. Hull and W. P. L. Mandy, *Am Acad Child Adolesc Psychiatry*, **56**, 466–474.
- 21 E. Bergers, G. Israel, C. Miller, B. Parkinson and N. Williams, *History 101: World History I (Malick and Gurian)*, [https://human.libretexts.org/Courses/Harrisburg\\_Area\\_Community\\_College/History\\_101%3A\\_World\\_History\\_I\\_\(Malick\\_and\\_Gurian\)/01%3A\\_Prehistory/1.05%3A\\_Agriculture\\_and\\_the\\_Neolithic\\_Revolution](https://human.libretexts.org/Courses/Harrisburg_Area_Community_College/History_101%3A_World_History_I_(Malick_and_Gurian)/01%3A_Prehistory/1.05%3A_Agriculture_and_the_Neolithic_Revolution).
- 22 A. E. Stark, *Twin Res Hum Genet*, **23**, 278–282.
- 23 M. J. Rantala, S. Luoto and J. I. Borráz-León, *Neuroscience Biobehavioral Reviews*, **122**, 28–37.