

Article

Extreme Tail Ratios and Overrepresentation among Subpopulations with Normal Distributions

Theodore P. Hill ^{1,*}  and Ronald F. Fox ^{2,†}¹ School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA² School of Physics, Georgia Institute of Technology, Atlanta, GA 30332, USA

* Correspondence: hill@math.gatech.edu

† These authors contributed equally to this work.

Abstract: Given several different populations, the relative proportions of each in the high (or low) end of the distribution of a given characteristic are often more important than the overall average values or standard deviations. In the case of two different normally-distributed random variables, as is shown here, one of the (right) tail ratios will not only eventually be greater than 1 from some point on, but will even become infinitely large. More generally, in every finite mixture of different normal distributions, there will always be exactly one of those distributions that is not only overrepresented in the right tail of the mixture but even completely overwhelms all other subpopulations in the rightmost tails. This property (and the analogous result for the left tails), although not unique to normal distributions, is not shared by other common continuous centrally symmetric unimodal distributions, such as Laplace, nor even by other bell-shaped distributions, such as Cauchy (Lorentz) distributions.

Keywords: tail ratio; normal distribution; overrepresentation; finite mixture of distributions



Citation: Hill, T.P.; Fox, R.F. Extreme Tail Ratios and Overrepresentation among Subpopulations with Normal Distributions. *Stats* **2022**, *5*, 977–984. <https://doi.org/10.3390/stats5040057>

Academic Editor: Wei Zhu

Received: 3 October 2022

Accepted: 18 October 2022

Published: 20 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In comparisons of the distributional values of a given psychological or physical trait between two populations (e.g., treated/untreated, male/female, exposed/nonexposed, elderly/youthful), the relative proportion of each population with values exceeding specified threshold or cutoff levels is often of more interest than comparisons of average values, standard deviations, or combinations thereof such as the Cohen *d* effect-size measure. The (right) tail ratio of one distribution compared to a second distribution, a measure of the relative tail proportions, is the ratio of the fraction of the first population above a given cutoff to the fraction of the second population above that same cutoff.

Tail ratios are a common measure of differences in extremes between populations in general and are of particular interest in psychological research, as emphasized by [1] in their review of the background, history, and practical advantages of tail ratios. For example, ref. [2] reported actual numerical ranges (between two and four) of certain male/female tail ratios and [3] studied how certain male/female tail ratios have changed over time. More generally, the issues of over- and underrepresentation of various factions of populations consisting of two or more subpopulations has become an important subject of study (see [4–8]).

The values of many biological or psychological traits (e.g., blood pressure, IQ, height) are often assumed to have normal (Gaussian) distributions (e.g., [9–11]), and the goal of this note is to record a simple fact about normal distributions that may be useful in interpreting statistical data concerning both tail ratios and over- and underrepresentation in mixed populations. In particular, it is shown below that in every population consisting of a finite number of subpopulations with different distributions of a normally distributed trait, exactly one of the subpopulations will not only dominate every other one in the right tail but also will do this in an extreme manner, eventually overwhelming all the other subpopulations. This property, although not unique to normal distributions, is not

shared by other common distributions, including ones that are also continuous, centrally symmetric, and unimodal, nor even by other bell-shaped distributions such as the common Cauchy (Lorentz) distributions.

2. Tail Ratios and Right-Tail Dominance

Recall that a probability measure P on the real line is uniquely determined by its complementary cumulative distribution function (ccdf) \bar{F}_P , defined by $\bar{F}_P(x) = P((x, \infty))$ for all $x \in \mathbb{R}$. (\bar{F}_P is also often called the survival function of P , since $\bar{F}_P(c)$ represents the P -probability of the set above the cutoff threshold c , i.e., the fraction that survives when all values less than or equal to c are removed.) With this notation, the formal definition of the tail ratio of P_1 to P_2 is as follows.

Definition 1. Given probability distributions P_1 and P_2 , the tail ratio of P_1 to P_2 is the function $\bar{F}_1(c)/\bar{F}_2(c)$, where \bar{F}_1 and \bar{F}_2 are the ccdfs for P_1 and P_2 , respectively.

(N.B. By convention, the tail ratio of P_1 to P_2 is that of the right tail; i.e., the ratio of the ccdf of P_1 to the ccdf of P_2 , not the ratio of the cdf of P_1 to the cdf of P_2 , which would yield the left tail ratio.)

A distribution P_1 may be said to dominate another distribution P_2 in the right tail if, for all sufficiently large cutoffs c , the tail ratio of P_1 to P_2 is strictly greater than 1, i.e., $\bar{F}_{P_1}(c) > \bar{F}_{P_2}(c)$. A much stronger notion of domination in the right tail is when the survival probabilities of P_1 eventually become arbitrarily larger than those of P_2 as the cutoff increases; this is formalized in the next definition.

Definition 2. A probability distribution P_1 strongly dominates distribution P_2 in the right tail if the (right) tail ratio of P_1 to P_2 becomes infinitely large as the cutoff increases, i.e., if

$$\lim_{c \rightarrow \infty} \frac{\bar{F}_{P_1}(c)}{\bar{F}_{P_2}(c)} = \infty.$$

We also recall that a continuous (absolutely continuous) probability distribution P is uniquely determined by its probability density function $f_P : \mathbb{R} \rightarrow [0, \infty)$ via $P((a, b)) = \int_a^b f_P(x)dx$; so, in particular, $\bar{F}_P(c) = \int_c^\infty f_P(x)dx$. The next lemma, which records a simple relationship between the quotients of probability density functions (pdfs) and the quotients of the corresponding ccdfs, is used in several examples and proofs below.

Lemma 1. Suppose P_1 and P_2 are continuous probability distributions with strictly positive continuous pdfs f_1 and f_2 and with ccdfs \bar{F}_1 and \bar{F}_2 , respectively. If $\lim_{x \rightarrow \infty} f_1(x)/f_2(x) = \alpha$, then $\lim_{c \rightarrow \infty} \bar{F}_1(c)/\bar{F}_2(c) = \alpha$.

Proof. Let f_1 and f_2 be strictly positive continuous pdfs with corresponding ccdfs \bar{F}_1 and \bar{F}_2 , respectively. Then, since $\lim_{x \rightarrow \infty} \bar{F}_1(x) = \lim_{x \rightarrow \infty} \bar{F}_2(x) = 0$, and $f_2(x) > 0$ for all x ,

$$\lim_{x \rightarrow \infty} \frac{\bar{F}_1(x)}{\bar{F}_2(x)} = \lim_{x \rightarrow \infty} \frac{\bar{F}'_1(x)}{\bar{F}'_2(x)} = \lim_{x \rightarrow \infty} \frac{f_1(x)}{f_2(x)} = \alpha,$$

where the first equality follows by the general form of L'Hôpital's rule, the second since $\bar{F}'_i(x) = -f_i(x)$, and the third by hypothesis. \square

The next example illustrates the difference between domination and strong domination.

Example 1. (i) Let P_1 and P_2 be Cauchy distributions with medians $m_1 = 0$ and $m_2 = 0.5$ and scale parameters $s_1 = 1$ and $s_2 = 0.5$, i.e., with density functions $f_{P_i}(x) = (\pi s_i [1 + ((x - m_i)/s_i)^2])^{-1}$, respectively. Then by Lemma 1,

$$\lim_{c \rightarrow \infty} \frac{\bar{F}_{P_1}(c)}{\bar{F}_{P_2}(c)} = \lim_{x \rightarrow \infty} \frac{f_{P_1}(x)}{f_{P_2}(x)} = \lim_{x \rightarrow \infty} \frac{\pi(1 + 2x^2 - 2x)}{\pi(1 + x^2)} = 2,$$

which implies that, as $c \rightarrow \infty$, the P_1 -probability of the set of numbers greater than c approaches exactly twice the P_2 -probability of numbers greater than c . Thus, although P_1 dominates P_2 in the right tail, neither P_1 nor P_2 strongly dominates the other in the right tail.

- (ii) Let P_1 and P_2 be Laplace distributions with medians $m_1 = 1$ and $m_2 = 0$ and scale parameters $s_1 = s_2 = 1$, i.e., with density functions $f_{P_i}(x) = (2s_i \exp(-(|x - m_i|)/s_i))^{-1}$, respectively. Then,

$$\lim_{c \rightarrow \infty} \frac{\bar{F}_{P_1}(c)}{\bar{F}_{P_2}(c)} = \lim_{c \rightarrow \infty} \frac{e^{1-c}}{e^{-c}} = e;$$

again, neither P_1 nor P_2 strongly dominates the other in the right tail.

- (iii) Let P_1 and P_2 be Laplace distributions with medians $m_1 = m_2 = 0$ and scale parameters $s_1 = 1$ and $s_2 = 0.5$, respectively. Then, P_1 strongly dominates P_2 in the right tail since

$$\lim_{c \rightarrow \infty} \frac{\bar{F}_{P_1}(c)}{\bar{F}_{P_2}(c)} = \lim_{c \rightarrow \infty} \frac{e^{-c}}{e^{-2c}} = \infty.$$

- (iv) Let P_1 and P_2 be normal distributions with identical variances $+1$ and with means 1 and 0 , respectively. Then, the density functions $f_{P_1}(x) = (1/\sqrt{2\pi})e^{-(1/2)(x-1)^2}$ and $f_{P_2}(x) = (1/\sqrt{2\pi})e^{-(1/2)x^2}$ for P_1 and P_2 , respectively, satisfy $(f_{P_1}(x))/(f_{P_2}(x)) = e^{x-(1/2)} \rightarrow \infty$ as $x \rightarrow \infty$; so, by Lemma 1, P_1 strongly dominates P_2 in the right tail.

3. Tail Ratios in Normal Distributions

When population research studies report only the means and standard deviations of their results, the default scientific understanding is that the data are approximately normally distributed. That is, the distributions in question are close to normal (Gaussian) distributions (see [12] for a comprehensive treatment of this classic distribution.)

For example, if a research study reports that their data have an average value of 2 and a standard deviation of 1, then the usual understanding is that the underlying dataset looks like the diagram in Figure 1(left) with $m = 2$ and $s = 1$, not like the somewhat similar Cauchy distribution in Figure 1(right). The underlying theoretical basis for the assumption of normality in most cases is the remarkable Central Limit Theorem, which says that if the numerical results of independent repetitions of *any experiment* are added, the empirical distribution (and consequently the sample average) always approaches a normal distribution. For instance, in the present context of tail ratios, the survey article by Voracek, Mohr, and Hagmann states “all tail-ratio calculations discussed here assume normally distributed variables” ([1], p. 882).

The appropriateness of assuming that given data have a normal distribution is often tested using the well-known empirical observation called the “68%–95%–99.7% rule” of normality illustrated in Figure 1(left). The one key property of a continuous centrally symmetric unimodal distribution that makes it normal is the unique (after rescaling) rate of decrease in its density function away from its mean. The normal density function, discovered by Gauss in 1809 in connection with his studies of astronomical observation errors, decreases from its mean at a rate exactly proportional to e^{-x^2} and not to e^{-x} or x^{-2} , for example, as is the case for the Laplace and Cauchy distributions, respectively.

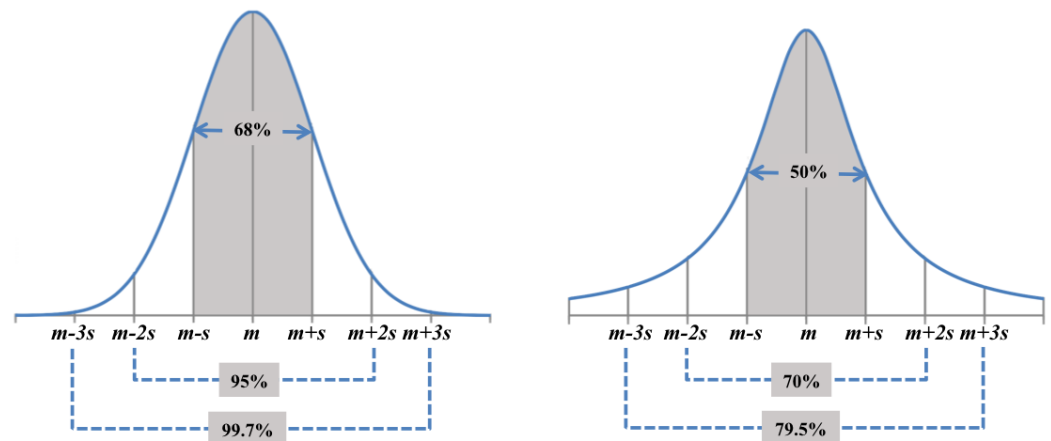


Figure 1. Universal empirical rules for all normal (left) and Cauchy (right) distributions. Here m represents the median, and s represents the distance (called standard deviation for normal distributions and scale parameter for Cauchy) from m to the inflection point. Although both families, normal and Cauchy, have similar bell-shaped density functions, normal distributions satisfy the main strong-dominance and overrepresentation properties presented in this note, but Cauchy distributions do not.

The Cauchy distribution, for instance, which sees widespread application in physics, also has a continuous centrally symmetric unimodal bell-shaped density similar to the normal distribution, but the Cauchy distribution has an undefined mean and variance and satisfies a different empirical rule, namely the 50%–70%–79.5% rule illustrated in Figure 1(right). Thus, it is very easy in practice to distinguish between these two similar-looking common bell-shaped distributions.

As is easy to see, the density functions of every two different normal distributions intersect in either exactly one or in exactly two distinct points. Thus, the density function of one of those two distributions is strictly larger than that of the other at all points greater than the larger of the two intersection points (or the unique one, if there is only one). This, in turn, implies that the proportion of that distribution from that point on is strictly larger than that of the other distribution from that point on; thus, this distribution will be overrepresented in the right tail. This is illustrated in the following example.

(N.B. For brevity, the standard notation $N(m, \sigma^2)$ will be used throughout this note to denote a normal distribution with mean m and standard deviation $\sigma > 0$.)

- Example 2.** (i) Let $P_1 \sim N(100, 10^2)$ and $P_2 \sim N(110, 10^2)$. It is clear that the unique crossing point of the density functions of P_1 and P_2 is at $x = 105$, which implies that the proportion of P_2 that is above any cutoff $c > 105$ is greater than the proportion of P_1 above c , i.e., the tail ratio of P_2 to P_1 is greater than 1 for all cutoff values c strictly greater than 105. Conversely, the proportion of P_1 below any $c < 105$ is greater than the proportion of P_2 below c .
- (ii) Let $P_1 \sim N(100, 10^2)$ and $P_2 \sim N(101, 11^2)$. By basic algebra, the two crossing points of the density functions of P_1 and P_2 are seen to be at $x_1 \cong 83.52$ and $x_2 \cong 106.95$, which implies that the tail ratio of P_2 to P_1 is greater than 1 for all cutoffs $c > x_2$. Similarly, in this case P_2 also dominates P_1 in the lower tail in that the proportion of P_2 that is below any cutoff $c < x_1$ is also greater than the proportion of P_1 below c .

As was seen in Example 1, for two given different Cauchy distributions or two different Laplace distributions, neither distribution may strongly dominate the other in the right tail. This is in sharp contrast to the main conclusion in this note, where it will be shown that in every finite collection of different normal distributions, there is always a unique one of those distributions that strongly dominates every one of the other distributions in the right tail.

Theorem 1. Let P_1 and P_2 be different normal distributions. Then,

- (i) either P_1 strongly dominates P_2 in the right tail or P_2 strongly dominates P_1 in the right tail;
- (ii) if P_1 strongly dominates P_2 in the right tail, then either P_1 has greater mean (average value) than P_2 or P_1 has greater variance than P_2 or both;
- (iii) if P_1 has greater variance than P_2 , then P_1 strongly dominates P_2 in both right and left tails, independent of the means.

Proof. Suppose that $P_1 \sim N(m, \sigma_1^2)$ and $P_2 \sim N(m, \sigma_2^2)$ are normal distributions with pdfs f_1 and f_2 , respectively. Since P_1 and P_2 are different, either $\sigma_1 \neq \sigma_2$ or $\sigma_1 = \sigma_2$ and $m_1 \neq m_2$.

Case 1. $\sigma_1 \neq \sigma_2$. Without loss of generality, $\sigma_1 > \sigma_2 = 1$. Then,

$$\begin{aligned} \frac{f_1(x)}{f_2(x)} &= \frac{\sqrt{2\pi}}{\sigma_1 \sqrt{2\pi}} \exp\left(\frac{-(x - m_1)^2}{2\sigma_1^2} + \frac{(x - m_2)^2}{2}\right) \\ &= \frac{1}{\sigma_1} \exp\left(\frac{x^2}{2} \left(1 - \frac{1}{\sigma_1^2}\right) + x \left(\frac{m_1}{\sigma_1^2} - m_2\right) + \left(\frac{m_1^2}{2} - \frac{m_2^2}{2}\right)\right). \end{aligned} \quad (1)$$

Since $\sigma_1^2 > 1$, $(1 - (1/\sigma_1^2)) > 0$, then (1) implies that $\lim_{x \rightarrow \infty} (f_1(x)/f_2(x)) = \infty$, and thus by Lemma 1, P_1 strongly dominates P_2 in the right tail.

Case 2. $\sigma_1^2 = \sigma_2^2$ and $m_1 \neq m_2$. Without loss of generality, $\sigma_1 = \sigma_2 = 1$. Then,

$$\begin{aligned} \frac{f_1(x)}{f_2(x)} &= \exp\left(\frac{-x^2 + 2m_1x - m_1^2 + x^2 - 2m_2x + m_2^2}{2}\right) \\ &= \exp\left((m_1 - m_2)x + \frac{(m_2^2 - m_1^2)}{2}\right). \end{aligned} \quad (2)$$

In addition, without loss of generality, $m_1 > m_2$, in which case (2) implies, via Lemma 1 as in Case 1, that P_1 strongly dominates P_2 in the right tail. This concludes the proof of (i); the proofs of (ii) and (iii) follow similarly. \square

The same essential argument extends easily to show that among every finite collection of different normal distributions, strong domination in the right tail by exactly one of those distributions is inevitable.

Corollary 1. Given a finite number of different normal distributions P_1, \dots, P_n , there is a unique one of these distributions that strongly dominates all the others in the right tail.

Proof. For each $i \in \{1, \dots, n\}$, let normal distributions P_i have mean m_i and standard deviation σ_i , respectively. Since the distributions are all different, if $m_i = m_j$ and $\sigma_i = \sigma_j$ then $i = j$, which implies that there exists a unique $i^* \in \{1, \dots, n\}$ such that $m_{i^*} = \max\{m_j : \sigma_j = \max\{\sigma_1, \dots, \sigma_n\}\}$. By the arguments for Cases 1 and 2 in Theorem 1, P_{i^*} strongly dominates P_i in the right tail for all $i \neq i^*$. \square

As was seen in Theorem 1, if P_1 has either greater variance than P_2 or the same variance and higher mean, then P_1 will strongly dominate P_2 in the right tail. Moreover, for many practical purposes, “most important, what might appear to be trivial group differences in both variability and central tendency can cumulate to yield very appreciable differences between the groups in numbers of extreme scorers” ([10], p. 11). The next example, a slight modification of the numerical example suggested by Feingold, illustrates this observation with the two normal distributions in Example 2(ii) that are close in mean value (100 vs. 101) and in standard deviation (10 vs. 11).

Example 3. Suppose a population X consists of two mutually exclusive subpopulations X_1 and X_2 , where the values of a given trait are normally distributed with distributions $P_1 \sim N(100, 10^2)$

and $P_2 \sim N(101, 11^2)$, respectively, as in Example 2(ii). A normal distribution calculator yields that $P(X_1 > 130) \cong 1.3490 \times 10^{-3}$ and $P(X_2 > 130) \cong 4.1900 \times 10^{-3}$, which yields the tail ratio $P(X_2 > 130)/P(X_1 > 130) \cong 3.106$. Thus, even with distributions this close in average value and standard deviation, if the two subpopulations X_1 and X_2 are of the same size, then the X_2 population will comprise more than 75% of the combined population beginning only three standard deviations above the mean.

Note that the results in Theorem 1 and Corollary 1 for normal distributions only depend on first-order asymptotic terms, and the question of which more general classes of distributions with rapidly decaying tails satisfy the same conclusions is left to the interested reader.

4. Overrepresentation in the Right Tail

Whether a particular subpopulation is overrepresented or underrepresented among the other subpopulations with respect to given values for a specific trait depends on the relative size of that subpopulation with those trait values compared to the size of the whole population with those trait values. For example, if subpopulation X_1 comprises 30% of the total population, but comprises 40% of the population with trait values above a given cutoff c , then X_1 is overrepresented in the portion of the total population with values greater than c .

The goal of this section is to show that a simple consequence of Corollary 1 is that in every finite mixture of different normal distributions, exactly one of those distributions will be strongly overrepresented in the right tail. (Recall that a *finite mixture of distributions* is a probability distribution with cdf F satisfying $F = \sum_{i=1}^n w_i F_i$, where $n > 1$, F_1, \dots, F_n are cdfs, and w_1, \dots, w_n are strictly positive weights with $\sum_{i=1}^n w_i = 1$.)

Definition 3. Given a finite mixture of distributions $F = \sum_{i=1}^n w_i F_i$, the distribution F_{i^*} is strongly overrepresented in the right tail of F if, as $c \rightarrow \infty$, the proportion of subpopulation F_{i^*} with values above c approaches 100% of the total population of F with values above c , that is, if

$$\lim_{c \rightarrow \infty} \frac{w_{i^*} \bar{F}_{i^*}(c)}{\sum_{i=1}^n w_i \bar{F}_i(c)} = 1.$$

Theorem 2. In every finite mixture of different normal distributions $F = \sum_{i=1}^n w_i F_i$, there is a unique $i^* \in \{1, \dots, n\}$ such that F_{i^*} is strongly overrepresented in the right tail of F .

Proof. For all $i \in \{1, \dots, n\}$, let \bar{F}_i denote the ccdf of P_i . By Corollary 1, there exists a unique $i^* \in \{1, \dots, n\}$ such that P_{i^*} strongly dominates P_i for all $i \neq i^*, i \in \{1, \dots, n\}$, i.e.,

$$\lim_{c \rightarrow \infty} \frac{\bar{F}_i(c)}{\bar{F}_{i^*}(c)} = 0 \quad \text{for all } i \neq i^*, i \in \{1, \dots, n\}. \tag{3}$$

Since $w_i > 0$ for all $i \in \{1, \dots, n\}$, (3) implies that

$$\lim_{c \rightarrow \infty} \frac{w_i \bar{F}_i(c)}{w_{i^*} \bar{F}_{i^*}(c)} = 0 \quad \text{for all } i \neq i^*, i \in \{1, \dots, n\},$$

which implies that

$$\lim_{c \rightarrow \infty} \frac{w_{i^*} \bar{F}_{i^*}(c)}{\sum_{i=1}^n w_i \bar{F}_i(c)} = 1,$$

so F_i^* is strongly overrepresented in the right tail of F . \square

Two concrete examples of normally-distributed traits among human populations are height and test scores such as those in the College Board Scholastic Aptitude Test (SAT). There are enormous amounts of data on human height, which are essentially continuous and are very close to being normally distributed ([11], p. 24). Scores on the SAT, on the other hand, are originally discrete but the distribution “obtained from a continuized, smoothed frequency distribution of original SAT scores” is a linear transformation of a normal distribution ([9], p. 59). Thus, since all linear transformations of normal distributions are normal, for all practical purposes, the resulting smoothed SAT scores have normal distributions.

Example 4. *The SAT scores of males and females are usually assumed (or designed) to be approximately normally distributed [9]. Unless the distributions are identical, Theorem 2 implies that exactly one of those two sexes must be strongly overrepresented in the right tail, and that this overrepresentation will increase as the score range increases; Figure 2 illustrates this with actual College Board statistics.*

New SAT Score Range	Male Students	Female Students	Male/Female Index (Raw)	Male Students	Female Students	Male/Female Index (Adjusted)
1400-1600	65,528	50,647	1.29	7.6%	5.2%	1.45
1200-1390	194,343	190,419	1.02	22.5%	19.7%	1.14
1000-1190	297,706	355,137	0.84	34.4%	36.7%	0.94
800-990	229,614	293,551	0.78	26.5%	30.3%	0.88
600-790	74,913	77,281	0.97	8.7%	8.0%	1.09
400-590	2,885	1,811	1.59	0.3%	0.2%	1.78
Total	864,989	968,846	0.89	100.0%	100.0%	1.00

Figure 2. Figure from [13], p. 7, titled “The numbers College Board didn’t publish” showing statistics for nearly two million students for the 2016 Edition of the Scholastic Aptitude Test, with breakdown by gender and score ranges. Note that the proportions of males in various score ranges, i.e., the tail ratios, increase as the score range increases and the left tail ratios also increase as the score range decreases. (About 10% more females participated than males, which is reflected in the Adjusted Male/Female Ratios).

5. Discussion

In real life examples, of course, there are no variables that are exactly normally distributed, since the normal distribution is continuous, and the number of people in various categories, for example, is necessarily finite. However, if distributions are close to being normally distributed, the right-tail overrepresentation of a unique subpopulation predicted by Theorem 2 (and the analogous conclusions for left tails) are perhaps reasonable to expect. Similarly, in real-life examples, calculations involving tails that are 6 or 7 standard deviations out involve probabilities of less than one in 10 billion and are meaningless among the current human population of this planet.

Author Contributions: Both authors contributed equally to this work and have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Voracek, M.; Mohr, E.; Hagmann, M. On the Importance of Tail Ratios for Psychological Science. *Psychol. Rep.* **2013**, *112*, 872–886. [[CrossRef](#)] [[PubMed](#)]
2. Hyde, J.S.; Mertz, J.E. Gender, culture, and mathematics performance. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 8801–8807. [[CrossRef](#)]
3. Lakin, J.M. Sex differences in reasoning abilities: Surprising evidence that male–female ratios in the tails of the quantitative reasoning distribution have increased. *Intelligence* **2013**, *41*, 263–274. [[CrossRef](#)]
4. Anand, R.; Winters, M.F. A Retrospective View of Corporate Diversity Training From 1964 to the Present. *Acad. Manag. Learn. Edu.* **2018**, *7*, 356–372. [[CrossRef](#)]
5. Ceci, S.J.; Williams, W.M.; Barnett, S.M. Women’s underrepresentation in science: Sociocultural and biological considerations. *Psychol. Bull.* **2009**, *135*, 218–261. [[CrossRef](#)] [[PubMed](#)]
6. Krebs, E.D.; Narahari, A.K.; Cook-Armstrong, I.O.; Chandrabhatla, A.S.; Mehaffey, J.H.; Upchurch, G.R.; Showalter, S.L. The Changing Face of Academic Surgery: Overrepresentation of Women among Surgeon-Scientists with R01 Funding. *J. Am. Coll. Surg.* **2020**, *231*, 427–433. [[CrossRef](#)]
7. Lett, E.; Murdock, H.M.; Orji, W.U.; Aysola, J.; Sebro, R. Trends in Racial/Ethnic Representation Among US Medical Students. *JAMA Netw. Open* **2019**, *2*, e1910490. [[CrossRef](#)] [[PubMed](#)]
8. Marum, R.J. Underrepresentation of the elderly in clinical trials, time for action. *Br. J. Clin. Pharmacol.* **2020**, *86*, 2014–2016. [[CrossRef](#)] [[PubMed](#)]
9. Dorans, N.J. Recentring and Realigning the SAT Score Distributions: How and Why. *J. Educ. Meas.* **2002**, *39*, 59–84. [[CrossRef](#)]
10. Feingold, A. The additive effects of differences in central tendency and variability are important in comparisons between groups. *Am. Psychol.* **1995**, *50*, 5–13. [[CrossRef](#)]
11. Roser, M.; Appel, C.; Ritchie, H. Human Height. 2019. Available online: <https://ourworldindata.org/human-height> (accessed on 30 September 2022).
12. Bryc, W. *The Normal Distribution: Characterizations with Applications*; Lecture Notes in Statistics; Springer: New York, NY, USA, 2012.
13. Sawyer, A. How the New SAT Has Disadvantaged Female Testers. 2017. Available online: <https://www.compassprep.com/new-sat-has-disadvantaged-female-testers/> (accessed on 30 September 2022).