Human Dimensions of Wildlife

An International Journal

ISSN: 1087-1209 (Print) 1533-158X (Online) Journal homepage: https://www.tandfonline.com/loi/uhdw20

Minimal effect sizes do not imply minimal effects for differences in long-tailed distributions

Jerry J. Vaske, Jay Beaman & Craig A. Miller

HUMAN DIMENSIONS OF WILDLIFE
2020, VOL. 25, NO. 3, 281–290
https://doi.org/10.1080/10871209.2020.1728790

Routledge
Taylor & Francis Group

Minimal effect sizes do not imply minimal effects for differences in long-tailed distributions

Jerry J. Vaske[a], Jay Beaman[a], and Craig A. Miller[b]

[a]Human Dimensions of Natural Resources, Colorado State University, Fort Collins, Colorado, USA; [b]Illinois Natural History Survey, University of Illinois, Champaign, Illinois, USA

ABSTRACT
Long-tailed distributions can distort findings, influence statistical tests, and result in small effect sizes. This research note proposed a definition of long-tailed distributions (i.e., $SD/M \geq 1$) and developed an alternative formulation of the Cohen's *d* effect size based on percent differences. Three hypotheses were examined: (a) waterfowl hunter harvest distributions tend to be long-tailed distributions, (b) differences in the means of two long-tailed distributions have minimal ($d < .2$) effect sizes unless the percent difference exceeds 20%, and (c) a minimal effect size does not necessarily imply that the difference in means should be ignored. Data obtained from 29 (1990–2018) annual waterfowl surveys in Illinois (n = 45,978) supported all three hypotheses. Statistical and managerial implications are discussed.

KEYWORDS
Long-tailed distributions; effect size; Cohen's *d*; duck harvest distributions

Introduction

Human dimensions variables of interest to wildlife managers (e.g., harvest estimates, days hunting, wildlife-related expenditures) often have long-tailed distributions (Moyer & Geissler, 1984, 1991; Vaske, Beaman, & Miller, 2020). A long-tailed harvest distribution arises when a large number of hunters have a low harvest rate, and a small number have a high harvest rate (Figure 1). Although the percentages of responses in the tail are small, these responses can inflate the standard deviation relative to the mean (Huan, Beaman, Chang, & Hsu, 2008). Long-tail distributions can distort findings, influence statistical significance tests, and result in small effect sizes (Miller & Anderson, 2002; Miller, Stephenson, & Williams, 2015; Vaske et al., 2020). This research note: (a) examined the prevalence of long-tailed distributions for Illinois mallard duck hunters from 1990 to 2018, (b) demonstrated that differences in the means of two long-tailed distributions had a minimal Cohen's *d* effect size unless the percent differences between the means was > 20%, and (c) illustrated that minimal effect sizes do not necessarily imply that the differences between the means of two long-tailed distributions do not have practical importance. In addressing these applied issues, the research note proposed a formal definition of long-tailed distributions based on the exponential distribution, and developed an alternative formulation of Cohen's *d* that is based on an intuitively meaningful percent differences in means, as opposed to the abstract traditional category cut points.

CONTACT Jerry J. Vaske jerryv@colostate.edu 244 Forestry, Human Dimensions of Natural Resources, Colorado State University, Fort Collins, CO 80523

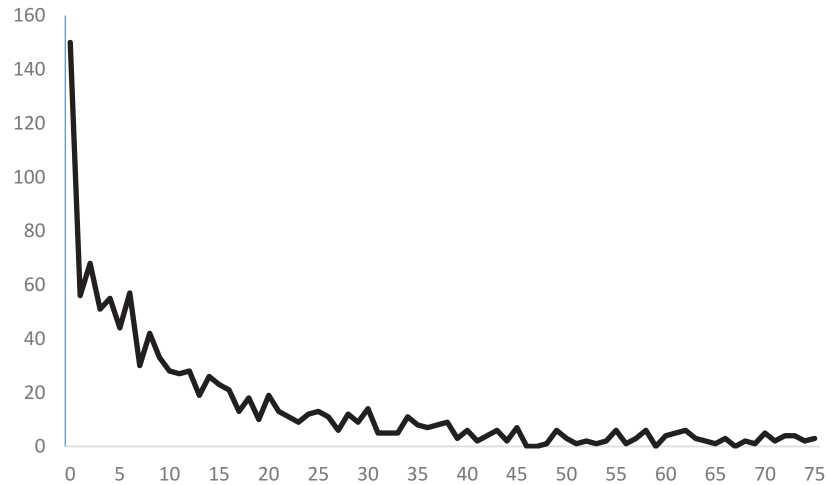

Figure 1. Example of a long-tailed distribution.

Long-Tailed Distributions

There are numerous examples of long-tailed distributions in the human dimensions of wildlife literature. Vaske et al. (2020), for example, examined the influence of record cards on the shape of response distributions from duck hunting surveys. Results indicated that individuals who received a record card were more likely to report small harvest rates and days hunting (i.e., beginning of the distribution), and less likely to give responses that contributed to heaping (i.e., middle of the distribution) compared to non-record card recipients. Record cards, however, did not influence the end of the distribution; both frequency functions were long-tailed distributions. In addition, effect sizes (e.g., point-biserial correlations [r_{pb}]) were less than .10, suggesting a minimal or small relationship (Kelly & Preacher, 2012; Vaske, 2019).

There is no formal definition of what constitutes a long-tailed distribution. Some researchers have suggested that when 80% of respondents contribute ≤ 50% to the mean, the distribution is long-tailed (HoggStuart & Klugman, 1983; Kreuter et al., 2014). Vaske et al. (2020), for example, used this definition and found that all percentages for duck hunters' annual harvest were consistently ≤ 50%. Among individuals who received preseason record cards for recording hunting activity, harvest percentages were 31% (mallard ducks), 28% (other duck species), and 36% (total ducks harvested). For non-record card respondents, the comparable percentages were 34%, 31%, and 39%, respectively. Based on responses to the 80th percentile accounting for 50% or less of the mean, all distributions were long-tailed.

Other researchers have made a connection between heavy-tailed distributions, long-tailed distributions, and the exponential distribution. For the exponential distribution, the mean (*M*) equals the standard deviation (*SD*) (Ross, 2014). A heavy-tailed distribution has a heavier tail than the exponential distribution since the exponential is a lower bound for a heavy-tailed distribution. By definition, heavy-tailed distributions extend from zero to plus and/or minus infinity (Foss, Korshunov, & Zachary, 2013).

Wildlife harvest distributions do not extend to infinity and do have an upper bound (e.g., the maximum number of birds legally harvested per individual hunter in a year under daily bag limits and a set season length), so mathematically such distributions cannot be heavy-tailed distributions. The same logic applies to days afield and expenditure distributions. There are true maximum numbers of days hunting per individual in a year and a maximum amount of money that each individual spent. Vaske (2019), for example, estimated the amount of money that survey respondents spent on all recreation equipment. On average, mail survey recipients spent \$4,569, whereas the average for telephone survey respondents was \$6,306. Despite a difference of \$1,737, the comparison of these means was not statistically significant due to large variances associated with the long-tailed distributions.

Distributions in wildlife harvest rates, hunter participation rates, and wildlife-related expenditures often appear to have exponential declines once the frequency of responses start to decline as response values increase. Given the exponential distribution: (a) is a lower bound for heavy-tailed distributions, and (b) has a *SD* that is equal to its *M* (Ross, 2014), a long-tailed distribution can be defined to occur when $SD/M \geq 1$. This definition applies to distributions with responses > 0. The definition also suggests that long-tailed distributions have tails heavier than exponential distributions in that $SD/M \geq 1$, not just $SD/M = 1$. Studies examining harvest distributions have consistently found the $SD/M \geq 1$ (e.g., Miller et al., 2015; Vaske et al., 2020).

Effect Sizes

An effect size is the strength of the relationship between an independent and a dependent variable (Coe, 2002; Kelly & Preacher, 2012). Effect size computations have been divided into two major types, often referred to as *r* family of indices and *d* family of indices (Rosenthal, 1994; Vaske, 2019). The *r* family of indices are expressed as correlation coefficients (Rosenthal, Rosnow, & Rubin, 2000). Using these indices, effect sizes are always ≤ 1.0, and range between − 1.0 and +1.0. The *d* family of indices are computed as the difference between the means of two groups (e.g., mean harvest for 2018 vs. 2019), divided by an appropriate standard deviation. Cohen's *d* (Cohen, 1988) is an example of this type of effect size:

$$\text{Cohen's } d = \frac{M_1 - M_2}{SD_p} \tag{1}$$

where SD_p is the pooled standard deviation.

A Cohen's *d* of .20 or less is commonly labeled a small relationship, .50 is a medium relationship, and .80 or more is a large relationship. Cohen (1988) provided examples of small, medium, and large effect sizes to support these suggested categories. Although a *d* of .80 or an *r* of .50 (which Cohen shows are mathematically equivalent) may not seem strong, Cohen (1988) argued that such values indicate " … grossly perceptible and therefore large differences," similar to "IQ difference between college graduates and persons with only a 50–50 chance of passing in an academic high school curriculum" (p. 27). Similarly, a correlation of .50 is "about as high as they come" in predictive effectiveness in applied psychology (Cohen, 1988, p. 81).

Cohen's (1988) labels (i.e., small, medium, large) were relative to each other and to the behavioral sciences in general. Such values, however, could vary among social science disciplines and specific research contexts (e.g., trends in wildlife harvest or hunter participation). For example, wildlife research has shown that the effect sizes associated with long-tailed distributions tend to be small (Miller et al., 2015; Vaske et al., 2020). The labeling is confounded because "small" effect sizes can sometimes have more practical importance than "large" effect sizes in other instances (Rosnow & Rosenthal, 1996). Sawilowsky (2009) added to the conventional categories by including "very small" and showing where this effect size can be important to recognize. Cohen (1988) argued that there is more to be gained than lost by offering a common frame of reference for evaluating effect size indices, especially when no better alternative exists for making a judgment. Vaske, Gliner, and Morgan (2002) agreed with Cohen's (1988) argument, but suggested a modification in the labels. For example, a "small effect" might be more appropriately labeled a "minimal relationship." A "medium effect" really reflects what is usual or "typical" for behavioral science studies and methods (i.e., a "typical relationship"). Similarly, a "large effect" might be categorized as a "substantial relationship." Substantial indicates that there is a clearly noticeable difference (*d*) or association (*r*). These modified labels clarify that the guidelines refer to a relationship, but not necessarily a relationship that has practical significance. The new labels also reinforce the relative nature of the relationships. A word such as "typical," for example, highlights that the relationship is common in the behavioral sciences.

In addition to the labeling issues, Cohen's *d* categories (i.e., .2, .5, .8) from Equation (1) are abstract and not necessarily intuitively meaningful. As noted earlier, the *d* family of effect size indices focus on changes or differences between means. Change and difference, however, can be computed in different ways. One approach focuses on temporal differences (e.g., year-to-year trends in mean harvest rates), whereas a second approach examines group or stakeholder differences (e.g., mean differences in hunter days afield by gender). Equation (2) gives a method for calculating a yearly trend. A researcher or manager, for example, may be interested in the percent change in harvest that occurred from year 1 to year 2, or if using record cards, reduced the bias by some percent (Vaske et al., 2020):

$$D = \frac{M_1 - M_2}{M_1} \tag{2}$$

where *D* stands for proportional difference.

Given Equations (1) and (2), Cohen's *d* can be re-written to include *D* and a variant of *SD/M*. In the ratio defining long-tailed distributions, *SD* and *M* are replaced with SD_p and M_1. SD_p in Equation (1) applies to the pooled standard deviation for two distributions (Equation (3)):

$$SD_p = \sqrt{\frac{n_1 SD_1^2 + n_2 SD_2^2}{n_1 + n_2}} \tag{3}$$

Equation (4) re-writes Cohen's *d* by introducing M_1/M_1 into Equation (1), essentially multiplying by 1. This results in the ratio $(M_1 - M_2)/M_1$ (i.e., *D*) and leaves M_1/SD_p as a multiplier (Equation (5)). Using $SD/M \geq 1$ in the definition of a long-tailed distribution emphasizes that the standard deviation exceeds the mean. Moving SD_p/M_1 to the denominator (Equation (6)) gives an alternative way of computing Cohen's d that provides an intuitive measure of change in relation to one mean (e.g., temporal change in year-to-year trends).

$$\text{Cohen's } d = \frac{M_1 - M_2}{SD_p}\left(\frac{M_1}{M_1}\right) \tag{4}$$

$$= \frac{M_1 - M_2}{M_1}\left(\frac{M_1}{SD_p}\right) \tag{5}$$

$$= \frac{D}{SD_p/M_1} \tag{6}$$

The above equations included M_1 in the denominator because this article is primarily concerned with temporal trends (e.g., year-to-year comparisons). To generalize these derivations to compare means between independent groups (e.g., males vs. females, rural vs. urban), the M_1/M_1 ratio can be replaced by the ratio of the average of two means (e.g., $\bar{U} = .5\ (\bar{U}_1 + \bar{U}_2)$), yielding Equation (7).

$$\text{Cohen's } d = \frac{D}{SD_p/\bar{U}} \tag{7}$$

Equations (6) and (7) convert the ratio in Equation (1) from an abstract statistic that has no intuitive meaning to a ratio that is meaningful to wildlife researchers and managers. The equation is not specific to long-tailed distribution and can be applied in any research. However, if two means differ by 10% and the distribution is long-tailed with $SD/M \geq 1$, Equation (6) predicts Cohen's $d = .1$ (i.e., less than minimal). Note that $d = .1$ does not depend on the value of M_1. If the percent change is 20% (i.e., a relatively large difference) and $SD_p/M_1 = 1$, Cohen's $d = .2$ (i.e., a minimal difference). If Cohen's $d = .2$ and $SD_p/M_1 = 1.5$, the percent change is 30%. This implies that a difference slightly < 30% can occur when $d < .2$. In other words, percent differences that are likely to have practical implications for planning and decision making can occur even when $d < .2$.

Based on this literature and these associated deductions, the following hypotheses were advanced:

H_1 Waterfowl hunter harvest distributions tend to be long-tailed distributions (i.e., $SD/M \geq 1$).

H_2 The difference in means of two long-tailed distributions has a minimal effect size unless the percent difference exceeds approximately 20%.

H_3 A minimal effect size for a difference in means of two long-tailed distributions does not necessarily imply that the difference can be dismissed.

Methods

The Illinois Natural History Survey has conducted annual waterfowl harvest surveys of Illinois hunters since 1981. Data for this research note were obtained from the 29 years of annual surveys conducted between 1990–1991 and 2017–2018. A random sample of state waterfowl stamp purchasers was selected each year. Following the conclusion of all Illinois waterfowl hunting seasons, participants were mailed a waterfowl hunter questionnaire. Four questionnaire mailings and postcard reminders were mailed in two-week intervals. Between 1990–1991 and 2017–2018, a total of 45,978 completed questionnaires were returned. Sample sizes for the selected study years ranged from 994 (in 2016) to 2,505 (in 2000). The average sample size was 1,585 with an average response rate of 68%.

Questionnaires contained items related to harvest. This article focused on the number of mallard ducks harvested. To address hypothesis one, the mean (*M*), standard deviation (*SD*), and *SD*/*M* were calculated for each study year. Hypothesis two was examined by computing independent samples *t*-tests, Cohen's *d*, the percent difference in the means, and approximated ±.05 confidence intervals for the percent difference. In these tests, the independent variables were the study years (e.g., 1990 vs. 1991, 1991 vs. 1992) and the dependent variable was the reported number of mallards harvested. Testing hypothesis three was based on the alternative formulation of Cohen's *d* (Equation (6)).

Results

Across all study years, the average number of mallards harvested was 9.24 per hunter per year. The means on a yearly basis ranged from 5.60 (in 1990–1991) to 12.35 (in 2000–2001) (Table 1). The standard deviations ranged from 7.23 (in 1990–1991) to 18.63 (in 2005–2006). All *SD*/*M* ratios were ≥ 1. The average ratio across all study years was 1.55. These ratios ranged from 1.09 (1991–1992) to 2.19 (1997–1998). Given that the definition of a long-tailed distribution is *SD*/*M* ≥ 1, these findings support hypothesis one.

For the 29 surveys, there were 28 year-by-year mean difference comparisons (e.g., 1990 vs. 1991, 1991 vs. 1992). Of these, the only *d* values >.2 were for percent differences greater than 20%. This is evidence for hypothesis two, but not proof. The proof stems from the algebra in the examples following Equation (7), which show that *d* cannot exceed .2 unless the percent difference > 20%. Therefore, hypothesis two is supported.

Hypothesis three predicted that a minimal effect size for a difference in means of two long-tailed distributions does not necessarily imply that the differences can be dismissed. Of the 23 comparisons with *d* ≤ .2 (Table 1), the percent differences ranged from 43.4% (1997–1998) to −2.0% (2009–2010). These values show that for *d* ≤ .2, some year-to-year percent differences occurred that should not be ignored. Six differences ranged from 20% to 43% and all of these differences were significant at $p < .001$. Logic suggests that these differences are not "minimal." In addition, six of the percent differences had absolute values exceeding 10% and only one of these was not statistically significant. These results support hypothesis three.

Discussion

Data from 29 annual surveys of Illinois waterfowl hunters consistently supported the first hypothesis that harvest distributions are long-tailed distributions (i.e., *SD*/*M* ≥ 1). These

Table 1. Long-tailed distributions for mallard harvest rates (1990–2018) in Illinois.

Year (Y)	Sample Size	*M*	*SD*	Y to Y Trend % *D*	±% for .05 CI trend	SD_p/M_1	Cohen's *d* Y to Y	*t*-value Y to Y *M* difference	*p*-value
1990	1385	5.60	7.23	35.90	10.7	1.46	0.25	6.55	< .001
1991	1430	7.61	8.98	−21.40	8.0	1.09	0.20	5.28	< .001
1992	1490	5.98	7.62	6.90	9.4	1.33	0.05	1.42	.154
1993	1562	6.39	8.30	−5.00	8.9	1.29	0.04	1.10	.269
1994	1676	6.07	8.20	69.00	13.5	1.90	0.36	10.02	< .001
1995	1528	10.26	14.34	−32.70	8.9	1.26	0.26	7.22	< .001
1996	1544	6.90	11.27	−8.60	12.4	1.62	0.05	1.35	.175
1997	1117	6.31	11.00	43.40	15.6	2.19	0.20	5.48	< .001
1998	1669	9.05	15.40	31.80	11.4	1.89	0.17	5.49	< .001
1999	2474	11.93	18.18	3.50	9.5	1.51	0.02	0.72	.466
2000	1582	12.35	17.89	−7.80	9.0	1.42	0.05	1.69	.089
2001	2505	11.39	17.33	−28.40	7.4	1.33	0.21	7.48	< .001
2002	2372	8.16	12.57	37.40	11.3	1.80	0.21	6.50	< .001
2003	1871	11.21	16.96	−25.00	8.9	1.40	0.18	5.50	< .001
2004	1980	8.41	14.45	31.70	12.9	1.97	0.16	4.83	< .001
2005	1738	11.08	18.63	9.30	11.0	1.67	0.06	1.65	.100
2006	1772	12.11	18.35	−11.50	10.4	1.47	0.08	2.17	.032
2007	1294	10.72	16.95	−7.50	11.8	1.56	0.05	1.24	.216
2008	1396	9.92	16.53	−8.70	12.3	1.65	0.05	1.38	.167
2009	1369	9.06	16.22	−2.00	12.9	1.72	0.01	0.30	.769
2010	1357	8.88	14.85	7.10	14.0	1.70	0.04	1.09	.276
2011	1374	9.51	15.30	18.40	11.6	1.72	0.11	2.57	.009
2012	1048	11.26	17.55	−8.30	10.9	1.49	0.06	1.40	.153
2013	1796	10.33	16.24	−13.10	14.8	1.54	0.09	2.34	.019
2014	1272	8.98	15.33	20.00	12.6	1.76	0.11	2.79	.007
2015	1154	10.78	16.34	−16.50	13.9	1.44	0.11	2.68	.011
2016	994	9.00	14.43	12.70	11.1	1.75	0.07	1.83	.076
2017	1520	10.14	16.51	−16.00	11.0	1.53	0.10	2.93	.004
2018	1709	8.52	14.64						

distributions that satisfy $SD/M \geq 1$ have inflated standard deviations relative to the mean. Of the year-to-year comparisons, only those with percent differences > 20% had Cohen's *d* effect sizes ≥ .2 (i.e., greater than minimal), supporting hypothesis two. Caution should be exercised when interpreting minimal effect sizes, however, because percent differences with absolute values of 20% or greater occurred for six minimal effect sizes and all these percent differences were significant. In addition, six percent differences for minimal effect sizes were over 10% and all but one of these was statistically significant. These findings supported hypothesis three; minimal effect sizes do not imply the year-to-year change in mean harvest differences can be dismissed as minimal.

Results here illustrated the importance of providing confidence intervals for estimates of trends. For example, a mean difference of −8.6% ± 12.4 for 1996 to 1997 could range from 4% to −21% and implies not using the −8.6% for planning. Similarly, the mean difference for 1997 to 1998 was 43.4% ±15.6% and this implies that the true estimate could range from 27.8% to 59.0%. For variables of interest to wildlife managers (e.g., harvest estimates, hunters' days afield, wildlife-related expenditures), researchers should provide not only effect size statistics, but also percent differences in means with their confidence intervals, as the latter typically have a more intuitive meaning when considering actionable suggestions.

Cohen's *d* is computed as the difference between two means, but the resulting categories of interpretation (i.e., .2, .5, .8) are of limited practical significance. This research note focused on alternative formulas for Cohen's *d* (Equations (6) and (7)) that convert

Cohen's *d* into an effect size that is easy to interpret (i.e., a percent difference in means divided by a standard deviation to mean ratio). Seven of the *d* values were ≥ .2 and all of these year-to-year trends had percent mean differences ≥ 20%. When *d* was < .2, there were three cases where the percent difference was ≥ 20% and six exceeded 10%. When expressed as a percent difference in means, Cohen's *d* has intuitive meaning; the *SD*/*M* ratio determines how a percent difference relates to *d*.

Calculating Cohen's *d* also allows computing a sample size that is necessary to detect a statistically significant difference with a given power (e.g., rejecting the null hypothesis at $p < .05$ with an 80% chance of detecting the difference). For example, to detect a 10% difference given $SD_p/M = 1$, Equation (6) gives $d = .1$. Using Soper's (2019) online sample size calculator (i.e., A-priori Sample Size Calculator for Student t-Tests), both means used for computing *d* should have 1,238 responses for a 1-tailed *t*-test. For a 2-tailed *t*-test, 1,571 responses are needed. This is consistent with Cohen's (1988) recommendations for avoiding Type II error.

Vaske et al. (2002) suggested that human dimensions researchers should address three questions when examining the relationships between variables. First, is an observed result real or should it be attributed to chance (i.e., statistical significance or *p*-values)? Second, if the finding is real, how large is it (i.e., effect size, Kim, 2015)? Third, is the effect large enough to be useful (i.e., practical significance, Kirk, 1996)? A significant *p*-value only indicates that there is some relationship or difference between the variables (i.e., the probability that an outcome could happen). Statistical significance, however, provides no information about the strength of the relationship (i.e., effect size) or whether the relationship has practical significance (Gliner, Vaske, & Morgan, 2001).

This article suggests that these questions need to be modified in a few ways when examining differences in means for long-tailed variables. First, rather than just considering statistical significance, Cohen's *d* should be used for determining the sample size necessary for detecting a significant difference in means (see Soper, 2019). Second, relying on effect size indicators is not sufficient for long-tailed distributions. For example, Cohen's *d* for 2011–2012 was .11 (i.e., less than minimal), yet the percent difference in means was 18.4% and the confidence interval was ± 11%. That is a difference that managers may not want to ignore. Third, if the magnitude of the relationship is a factor of importance, effect size indices can be good indicators of practical significance (Grissom & Kim, 2005). For *SD*/*M* ≥ 1 variables, however, practical significance needs to be taken in the context of Equations (6) and (7) for Cohen's *d*.

Not all similar effect sizes are equally important. A "small" effect size may have more practical significance than a "large" effect size in one instance, but not in another. Rosnow and Rosenthal (1996), for example, examined Aspirin's effect on heart attacks and demonstrated that those who took Aspirin had a statistically significant lower probability of having a heart attack than those in the placebo condition. Although the effect size for this relationship was minimal, the practical importance was high, due to both the low cost of taking Aspirin and the importance of reducing myocardial infarction. If this same effect size were found for a prescription for a more expensive heart medication, the practical significance might be lower.

Overall, Equations (6) and (7) expressed Cohen's *d* as a percent difference divided by SD_p/M_1. These equations apply to any difference of means. There are, however, some unresolved questions. First, research is needed on what variables of interest to wildlife managers (e.g., hunter participation, expenditure data) tend to be long-tailed ($SD/M \geq 1$). Second, if $SD/M \geq 1$ variables are common for temporal differences (e.g., trend analysis) such as those examined here, research exploring the general implications of such variations is warranted. Third, this research note examined year-to-year comparisons to illustrate the approach. It would be valuable to explore trends across longer time periods (e.g., decades). Fourth, the daily bag limits and season lengths in Illinois have remained unchanged since the mid-1990s, and the shapes of the distributions for each annual harvest remained relatively constant. Not explored, however, are the potential consequences of new polices or changes in policies that could alter the shape of the long-tailed distributions. Fifth, research applying the percent difference approach to mean comparisons between independent groups (e.g., males vs. females, rural vs. urban) would further the generalizability of the findings. Finally, a long-tailed distribution was defined as $SD/M \geq 1$. Results supported this definition, but given the lack of an accepted definition, research exploring the generalizability of the definition would be prudent.

References

Coe, R. (2002, September 12–14). *It's the effect size, stupid: What effect size is and why it is important*. Paper presented at the annual conference of the British educational research association. Exeter, UK: University of Exeter. Retrieved from www.leeds.ac.uk/educol/documents/00002182.htm

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Foss, S. G., Korshunov, D., & Zachary, S. (2013). An introduction to heavy-tailed and subexponential distributions (Chapter 2). In Springer Series in *Operations research and financial engineering, science and business media,* chapter 2, p. 7. New York, NY: Springer.

Gliner, J. A., Vaske, J. J., & Morgan, G. A. (2001). Null hypothesis significance testing: Effect size matters. *Human Dimensions of Wildlife*, *6*, 291–301. doi:10.1080/108712001753473966

Grissom, R. J., & Kim, J. J. (2005). *Effect sizes for research: A broad practical approach*. Mahwah, NJ: Erlbaum.

HoggStuart, R., & Klugman, A. (1983). On the estimation of long-tailed skewed distributions with actuarial applications. *Journal of Econometrics*, *23*, 91–102. doi:10.1016/0304-4076(83)90077-5

Huan, T. C., Beaman, J., Chang, L. H., & Hsu, S. Y. (2008). Robust and alternative estimators for "better" estimates for expenditures and other "long tail" distributions. *Tourism Management*, *29*, 795–806. doi:10.1016/j.tourman.2007.09.004

Kelly, K., & Preacher, K. (2012). On effect size. *Psychological Methods*, *17*, 137–152. doi:10.1037/a0028086

Kim, H. Y. (2015). Statistical notes for clinical researchers: Effect size. *Restorative Dentistry & Endodontics*, *40*, 328–331. doi:10.5395/rde.2015.40.4.328

Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, *56*, 746–759. doi:10.1177/0013164496056005002

Kreuter, M. W., Hovmand, P., Pfeiffer, D., Fairchild, M., Rath, S., Golla, B., & Casey, C. (2014). The "long tail" and public health: New thinking for addressing health disparities. *American Journal of Public Health*, *104*, 2271–2278. doi:10.2105/AJPH.2014.302039

Miller, C. A., & Anderson, W. L. (2002). Digit preference in reported harvest among Illinois waterfowl hunters. *Human Dimensions of Wildlife*, *7*, 55–65. doi:10.1080/108712002753574783

Miller, C. A., Stephenson, A. L., & Williams, B. D. (2015). Reported harvest and days afield among waterfowl hunters: Do pre-season contacts make a difference? *Human Dimensions of Wildlife*, *20*, 182–184. doi:10.1080/10871209.2014.971473

Moyer, L. M., & Geissler, P. H. (1984). Improving stability of annual state waterfowl harvest estimates in highly-skewed data. *American Statistical Association, Proceedings of the section on survey research methods* (pp. 467–471), Philadelphia, PA.

Moyer, L. M., & Geissler, P. H. (1991). Accommodating outliers in wildlife surveys. *Wildlife Society Bulletin*, *19*, 267–277.

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. Hedges (Eds.), *The handbook of research synthesis* (pp. 231–244). New York, NY: Russell Sage Foundation.

Rosenthal, R., Rosnow, R. L., & Rubin, D. B. (2000). *Contrasts and effect sizes in behavioral research: A correlational approach*. Cambridge, UK: Cambridge University Press.

Rosnow, R. L., & Rosenthal, R. (1996). *Beginning behavioral research* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Ross, S. M. (2014). *Introduction to probability and statistics for engineers and scientists* (5th ed.). New York, NY: Elsevier.

Sawilowsky, S. S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*, *8*, 597–599. doi:10.22237/jmasm/1257035100

Soper, D. S. (2019). A-priori sample size calculator for student t-tests [Software]. Retrieved from http://www.danielsoper.com/statcalc

Vaske, J. J. (2019). *Survey research and analysis* (2nd ed.). Urbana, IL: Sagamore-Venture.

Vaske, J. J., Beaman, J., & Miller, C. A. (2020). Influence of advanced notice record cards on waterfowl hunting participation and harvest distributions. *Human Dimensions of Wildlife*, *25*(3), this issue.

Vaske, J. J., Gliner, J. A., & Morgan, G. A. (2002). Communicating judgments about practical significance: Effect size, confidence intervals and odds ratios. *Human Dimensions of Wildlife*, *7*, 287–300. doi:10.1080/10871200214752