Teaching and Education Commentary

Impossible Inferences: A Note on Inconsistency in Business Statistics

Joseph G. Eisenhauer[a]
[a]*University of Detroit Mercy*

JEL Codes: A22, C12, I21, M21
Keywords: confidence interval, descriptive statistics, hypothesis testing, inference, population standard deviation, *Z* statistic

Abstract
Textbooks designed for business statistics customarily present inferential statistics whose premises contradict the definitions developed in descriptive statistics. In particular, because the standard deviation of a population cannot be calculated if the population mean is unknown, hypothesis tests and confidence intervals for the mean that rely on the population standard deviation pose an inconsistency between lessons. To reduce potential confusion and utilize both class time and students' homework time efficiently, the present note proposes de-emphasizing these tests and intervals in favor of those that depend on the sample standard deviation.

1 Introduction

Because the analysis of empirical data is central to good business decision-making, most business curricula require a course in business statistics (Mitra 2022). To the extent that such a course is distinct from statistics courses taught elsewhere in the university, business statistics is ideally somewhat less theoretical, with more applications to realistic problems that could be encountered in practice (Liu 2017). Indeed, students who select the business major are typically pragmatists, strongly motivated by career opportunities (Kim, Markham, and Cangelosi 2002; Granitz, Chen, and Kohli 2014; Matthews, Edmondson, and Matthews 2024), who respond well to courses that mimic job responsibilities (Debnath, Tandon, and Pointer 2007). To maintain students' interest, Simons (2020) introduces a semester-long exercise analyzing real-world data from a field experiment.

Meanwhile, the knowledge base of statistics, like every academic discipline, naturally expands over time. To keep up with this growth, introductory-level textbooks tend to add new material (including online appendices) without eliminating much, becoming increasingly encyclopedic, lengthier, and—to the chagrin of many students—more expensive.[1] Instructors, meanwhile, face the ever-more daunting task of selecting what subject matter to include, and what to exclude, in a finite period of time—a year, a semester, or a quarter. The pedagogical journal literature often exacerbates the problem, by proposing new topics that could be presented. It is, of course, nice to have options, but every such choice comes with the opportunity cost of forgoing a different alternative. Rarely are suggestions offered for reducing the canon of introductory statistics.

The present note is an exception. It suggests that the familiar hypothesis test for the mean of a single population when its standard deviation is known, the related test for the difference in the means of two populations when their variances are known, and the corresponding confidence intervals are inconsistent with descriptive statistics and are thus impractical undertakings that not only can, but

[1] Textbooks are almost always the primary curricular material in such courses (Harter and Asarta 2022).

should, be minimized in a business statistics course. This would make the flow of the course more sensible and save time for more valuable content.

2 A Logical Fallacy

Prior to studying statistical inference, students have inevitably learned that the population standard deviation is calculated as

$$\sigma = \sqrt{\frac{\sum(x_i - \mu)^2}{N}}, \tag{1}$$

where X is the variable of interest, μ is the population mean, and N is the population size. At this point, students have become aware that μ must be known in order to calculate σ. Yet, the first hypothesis test that students encounter is generally the large sample test for a single population mean when σ is known. The test statistic is given as

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}, \tag{2}$$

where $\bar{x}$ is the sample mean, n is the sample size, and μ_0 is the hypothesized value of the population mean (Peck, Olson, and Devore 2015; Moore et al. 2016; Anderson et al. 2024; Lind, Marchal, and Wathen 2024). The rationale for the use of this statistic is the Central Limit Theorem, which states that for large samples, $\bar{x}$ is normally distributed with a mean equal to μ and a standard error equal to $\sigma/\sqrt{n}$. The presentation of the test statistic is commonly followed by numerical examples and several practice problems to be assigned as homework.

But there is a logical error here: if the true population mean, μ, is unknown (which is, of course, the reason for conducting the test), then the population standard deviation, σ, cannot be known, and thus cannot be used as in Equation (2). Teaching students to conduct such a test, complete with numerical illustrations and homework assignments, creates a contradiction in lessons that can be—indeed, ought to be—rather confusing for students who have been paying close attention throughout the course.[2]

The examples given to illustrate the use of this test are necessarily implausible. In some instances, there is simply an assertion that the investigator, as if by magic, somehow divines σ without knowing μ. Other authors attempt to explain how σ could be known, by conflating the population of interest with some other grouping. Consider the following example (Moore et al. 2016): a firm sells 473-milliliter (ml) bottles of iced tea. Prior experience shows that the population standard deviation is 2 ml. After a new filling process is installed, the firm samples 20 bottles and finds a sample mean of 472.56 ml (but does not bother to calculate the sample standard deviation). To test the null hypothesis, H_0: $\mu = 473$, against the alternative, H_a: $\mu \neq 473$, the test statistic is calculated as $z = (472.56 - 473)/(2/\sqrt{20}) = -0.984$, which gives a prob-value of 0.328, so the null is not rejected. Aside from the obvious error of using a large-sample statistic (Z) with a small sample, the standard deviation is incorrect. The population of

[2] Such issues may contribute to the perception of statistics as a difficult and confusing course. Students are especially confused by statistical inference (Soto et al. 2007; Dambolena, Eriksen, and Kopcso 2009). Stevens and Palocsay (2012, p. 132) found that "inferential concepts and calculations were considerably more difficult for students than was rote calculation," with only 69 percent able to compute the standard error of the mean in an inferential problem. Mitra (2023) found that 15 percent of students who had studied business statistics were dissatisfied with the course, 20 percent lacked confidence in their understanding of course material, and more than 11 percent lacked confidence in their ability to identify the appropriate statistical methods for solving problems. Both Zanakis and Valenzi (1997) and Sharma and Srivastav (2021) discovered that students actually had less interest in statistics, and considered it less worthwhile, after completing the business statistics course.

interest consists of bottles filled by the new process, not those filled by the historical process. If the new process has potentially changed the mean, it is likely to have changed the standard deviation also. At best, the old σ is a proxy for the new σ, and it may or may not be a better proxy than the sample standard deviation that could easily be computed from the actual observations. (Indeed, it is questionable whether the investigator could even have known that the old population had a standard deviation of 2, without having inspected every single bottle.)

Similar problems occur in many popular texts (Peck, Olson, and Devore 2015; Anderson et al. 2024; Lind, Marchal, and Wathen 2024). Students find such artificially convenient examples perplexing and unmotivating. As Shiu (2013, p. 13) notes, "A major potential problem of making up these questions and data is that students may not feel their relevance to the real world, and feel the learning process of dealing with these questions and data uninspiring. These feelings inevitably exert an adverse impact on the learning outcome."

Still other textbooks explicitly impose strong assumptions to justify the use of σ when μ is unknown. In comparing a treatment group to a control group, for example, Gravetter and Wallnau (2017, p. 226) state, "To simplify the hypothesis-testing situation, one basic assumption is made about the effect of the treatment: if the treatment has any effect, it is simply to add a constant amount to (or subtract a constant amount from) each individual's score." This clever trick implies that the treatment can shift the mean without altering the standard deviation, though the realism of that assumption is never addressed. A number of worked problems and exercises follow that depend on knowing σ. To their credit, the authors acknowledge, albeit not until a subsequent chapter, "In most situations, however, the standard deviation for the population is not known. In fact, the whole reason for conducting a hypothesis test is to gain knowledge about an *unknown* population" (Gravetter and Wallnau 2017, p. 269, emphasis inoriginal).

Indeed, all textbooks rightly instruct students to replace σ with the sample standard deviation $s = \sqrt{\sum(x_i - \bar{x})^2/(n-1)}$ when σ is unknown and calculate the large-sample test statistic $Z = (\bar{x} - \mu_0)/\left(s/\sqrt{n}\right)$ or its small-sample counterpart, $t = (\bar{x} - \mu_0)/\left(s/\sqrt{n}\right)$. In practice, this should always be done, because σ cannot be measured if μ is unmeasured. A test that assumes knowledge of σ without knowledge of μ is logically inconsistent and uses up scarce class time.

3 Other Impossible Inferences

Similar to the one-population hypothesis test, textbooks customarily present the initial test for the difference between two population means using

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \tag{3}$$

as the test statistic, where D_0 is the hypothesized difference between the population means and σ_i^2 is the ith population variance. But here again, in order to compute the population variances, the two population means would already have to have been calculated, in which case there is no need to hypothesize about the magnitude of their difference. This test, too, should begin from the premise that neither population variance is known and substitute the sample variances in their places.

Finally, the same false reasoning occurs when, instead of hypothesis testing, a confidence interval for the population mean is constructed as $\bar{x} \pm z_{\alpha/2}\left(\sigma/\sqrt{n}\right)$, where α is the level of significance and $z_{\alpha/2}$ is the value of the standard normal variable with an area of $\alpha/2$ in each tail, or when a confidence interval for the difference in two means is constructed as $(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}$. Regardless of whether the inference takes the form of a confidence interval or hypothesis test, the concept of a known σ and unknown μ is problematic.

4 Conclusion

The Central Limit Theorem assumes that the population parameters are known and asks how the means of large samples would be distributed around μ; fundamentally, it is an exercise in statistical deduction, arguing from the general to the specific. By contrast, hypothesis testing is inherently an exercise in statistical induction, or inference, arguing from the specific to the general when the population parameters are unknown. While the Central Limit Theorem provides a useful insight into the construction of the Z statistic, the Z test with known σ is an awkward amalgamation that paradoxically purports to investigate an unknown population by relying on knowledge of the very population being studied. Thus, instructors struggling to fit the enormous volume of material from introductory textbooks into a limited amount of course time, and those seeking greater consistency across lessons, may want to consider streamlining the presentation by only briefly mentioning the Z statistic with known σ as a purely theoretical construct, and immediately replacing it with a statistic that relies on the sample standard deviation. Such an approach has been used in textbooks (McClave, Benson, and Sincich 2016), but rarely so. The more common practice of presenting Z with a known σ, offering rather strained examples to illustrate it, and asking students to employ it to solve numerical problems, represents a potentially confusing expenditure of time and effort that could surely be better spent elsewhere.[3]

About the Authors: Joseph G. Eisenhauer is a Professor of Economics and Dean of the College of Business Administration at the University of Detroit Mercy (Email : eisenhjg@udmercy.edu).

[3] Anecdotal evidence suggests that some instructors who thoroughly teach the Z statistics with known σ—discussing the concepts and calculations in class, providing numerical examples, assigning related homework, and including these on quizzes or tests—are nonetheless uncomfortable with this approach. An instructor at an Agricultural and Applied Economics Association (AAEA) member institution acknowledged using this approach but noted, "My opinion is that the population variance is almost never (if ever!) known, and a modern statistics course might be better served focusing efforts on teaching methods more relevant to the kinds of data students will encounter in the modern world." Others adopt the abbreviated approach recommended here, despite the textbook treatment. As a referee for this journal mentioned, "I… and many of my colleagues at different universities… just mention the Z statistics because it is not really practical to teach [them], for the very reasons that are outlined in this paper."

References

Anderson, D.R., D.J. Sweeney, T.A. Williams, J.D. Camm, J.J. Cochran, M.J. Fry, and J.W. Ohlmann. 2024. *Statistics for Business & Economics*, 15th ed. Boston, MA: Cengage.

Dambolena, I.G., S.E. Eriksen, and D.P. Kopcso. 2009. "An Intuitive Introduction to Hypothesis Testing." *INFORMS Transactions on Education* 9(2):53–62.

Debnath, S.C., S. Tandon, and L.V. Pointer. 2007. "Designing Business School Courses to Promote Student Motivation: An Application of the Job Characteristics Model." *Journal of Management Education* 31(6):812–831.

Granitz, N., S. Chen, and K.K. Kohli. 2014. "Choosing Business as a College Major: A Survey of High School Students." *Journal of the Academy of Business Education* 15:1–22.

Gravetter, F.J., and L.B. Wallnau. 2017. *Statistics for the Behavioral Sciences*, 10th ed. Boston MA: Cengage.

Harter, C., and C.J. Asarta. 2022. "Teaching Methods in Undergraduate Intermediate Theory, Statistics and Econometrics, and Other Upper-Division Economics Courses: Results from a Sixth National Quinquennial Survey." *American Economist* 67(1):132–146.

Kim, D.F., S. Markham, and J.D. Cangelosi. 2002. "Why Students Pursue the Business Degree: A Comparison of Business Majors Across Universities." *Journal of Education for Business* 78(1):28–32.

Lind, D.A., W.G. Marchal, and S.A. Wathen. 2024. *Statistical Techniques in Business and Economics*, 19th ed. New York: McGraw-Hill Irwin.

Liu, Z. 2017. "Teaching Reform of Business Statistics in College and University." *Eurasia Journal of Mathematics, Science and Technology Education* 13(10):6901–6907.

Matthews, K.N., D.R. Edmondson, and L. Matthews. 2024. "Why Students Select Their College Major: An Investigative Study." *Atlantic Marketing Journal* 13(1):Article 11.

McClave, J.T., P.G. Benson, and T. Sincich. 2016. *Statistics for Business and Economics*, 13th ed. London: Pearson.

Mitra, S. 2022. "Teaching Statistics and Data Science to Business Students." *Chance* 25(4):27–35.

———. 2023. "How Are Students Learning in a Business Statistics Course? Evidence from Both Direct and Indirect Assessment." *INFORMS Transactions on Education* 23(2):95–103.

Moore, D.S., G.P. McCabe, L.C. Alwan, and B.A. Craig. 2016. *The Practice of Statistics for Business and Economics*, 4th ed. New York: W.H. Freeman.

Peck, R., C. Olson, and J.J. Devore. 2015. *Introduction to Statistics and Data Analysis*, 5th ed. Pacific Grove CA: Brooks/Cole.

Sharma, A.M., and A. Srivastav. 2021. "Study to Assess Attitudes Toward Statistics of Business School Students: An Application of the SATS-36 in India." *International Journal of Instruction* 14(3):207–222.

Shiu, E. 2013. "Improving Learning in Statistics—A Pilot Trial Study to Achieve the Triple Effects of Depth, Breadth and Integration." *International Journal of Management Education* 11(1):12–24.

Simons, A.M. 2020. "Making Business Statistics Come Alive: Incorporating Field Trial Data from a Cookstove Study into the Classroom." *Applied Economics Teaching Resources* 2(3):76–86.

Soto, A.E.C., S. Vanhoof, W. Van den Noortgate, and P. Onghena. 2007. "Students' Misconceptions of Statistical Inference: A Review of the Empirical Evidence from Research on Statistics Education." *Educational Research Review* 2(2):98–113.

Stevens, S.P., and S.W. Palocsay. 2012. "Identifying Addressable Impediments to Student Learning in an Introductory Statistics Course." *INFORMS Transactions on Education* 12(3):124–139.

Zanakis, S.H., and E.R. Valenzi. 1997. “Student Anxiety and Attitudes in Business Statistics.” *Journal of Education for Business* 73(1):10–16.

DOI: https://doi.org/10.71162/aetr.224304.