EFFECT OF CALCULATOR TECHNOLOGY ON STUDENT ACHIEVEMENT IN AN INTRODUCTORY STATISTICS COURSE

LINDA BRANT COLLINS
University of Chicago
collins@galton.uchicago.edu

KATHLEEN CAGE MITTAG
University of Texas at San Antonio
kmittag@utsa.edu

SUMMARY

We report on a study of the relationship between calculator technology and student learning in two introductory statistics class sections taught by the same instructor at the University of Texas at San Antonio. At the introduction of hypothesis testing and confidence intervals, one class section (A) was given graphing calculators capable of inferential statistics to use for a few weeks. At the same time, the other class section (B) was given non-inferential graphing calculators. Data were collected on all test grades and daily quiz grades for both class sections. The students were allowed to use the inferential calculators on only the examination covering hypothesis tests and confidence intervals and on the final examination. Both sections received the same tests. We found that although use of the calculator with inferential capabilities is associated with improved scores on the inferential examination, the improvement is not significant once we adjust for performance on previous tests. Still, we note that on final examination questions specifically utilizing the calculator inference functions, the two classes perform similarly. In fact, both classes had trouble with "calculations" while at the same time answering "concept" questions fairly well. The inferential calculator did not appear to give students any clear advantage or disadvantage in their performance on examinations.

Keywords: Statistics education research; Introductory statistics; Graphing calculator; Inferential calculator; Student achievement

1. INTRODUCTION

Since calculators with inferential statistics capabilities came on the market in January, 1996, it has become evident that statistics educators need to analyze the effectiveness of the new hand-held technology. It is interesting to note that many statistics instructors at our university were not aware of these calculators and this may also be true at many other universities.

We studied the effect of calculator technology on student achievement in two introductory statistics class sections taught by one of the authors in autumn, 1998, at a large public urban university in the United States. At the introduction of the topics of hypothesis testing and confidence intervals, one class section (Class A) was given inferential calculators to use for a few weeks. At the same time, the other class section (Class B) was given older calculators without inferential capabilities. Other than this difference in calculators, the two groups were treated as similarly as possible.

Statistics Education Research Journal, 4(1), 7-15, http://www.stat.auckland.ac.nz/serj

In this paper, we first review the current literature on technology in the statistics (and mathematics) classroom and then proceed to an analysis of the data collected from our own study.

2. BACKGROUND: COMPUTERS

Many efforts are being made to enhance the learning experiences for students in introductory statistics courses (Cobb, 1993; Garfield, 1995; Gnanadesikan, Scheaffer, Watkins, & Witmer, 1997). Technology is influencing the teaching and learning of statistics. Gilchrist (1986) suggested that computers should be utilized to teach concepts and methods and that number crunching should be de-emphasized. Singer and Willet (1990) asserted that since the advent of computers, artificial data sets are no longer needed. David Moore (1992) suggested that calculation and graphics be automated as much as possible. Hogg (1992), Neter (1989) and Pirie (1989) wrote that the use of computers in undergraduate statistics classrooms is very important. A computer-based instructional strategy should be used either for managing large data sets or for generating simulations to illustrate probability concepts (Mittag & Eltinge, 1998).

Moore wrote in Moore, Cobb, Garfield, and Meeker (1995) that

> The nature of both statistical research and statistical practice has changed dramatically under the impact of technology. Our teaching has certainly changed as well --- but what strikes me is how little it has changed. The computing revolution has changed neither the nature of teaching nor our productivity as teachers. (p. 250)

Moore goes on to suggest that some reasons for educators being slow to change may be: “our costs have risen much faster than incomes or inflation; we see no need to change; we have an outdated organizational structure; and there is little internal incentive to change” (p. 251).

3. BACKGROUND: CALCULATORS

The advent of calculator technology has influenced the teaching of mathematics in a profound way (Dunham & Dick, 1994; Demana & Waits, 1990; Fey & Good, 1985). Several research studies have documented the benefits of calculator use in the mathematics classroom (Campbell & Stewart, 1993; Carlson, 1995; Dunham, 1993; Dunham, 1996; Graham & Thomas, 2000; Harvey, Waits, & Demana, 1995; Hembree & Dessart, 1986; Hennessy, Fung, & Scanlon, 2001; Quesada & Maxwell, 1994). Research has not focused on the use of calculator technology in the statistics classroom. There have been a few papers published which discuss graphing calculators in statistics. Rinaman (1998) discussed changes that had been made in basic statistics courses at his university. The TI-83 graphing calculator was recommended for the students to purchase but it was determined that all students should be required to purchase this or no calculator should be allowed at all. Garfield (1995, p.31) wrote that “calculators and computers should be used to help students visualize and explore data, not just to follow algorithms to predetermined ends.” A sample lesson using the TI-80 calculator based on modeling and simulation was discussed by Graham (1996a). Binomial graph and Poisson graph programs for the TI-82 were presented and demonstrated by Francis (1997). Statistical features of the TI-83 and TI-92 were discussed by Graham (1996b). Since there has been little, if any, research on the effect of graphing calculators on conceptual understanding in introductory statistics, the authors decided to conduct the study described in this paper.

4. EXPLANATION OF CALCULATORS USED IN THE STUDY

Statistics students have used the scientific calculator for the past two decades. With the introduction of the graphing calculator about ten years ago, basic descriptive statistics and graphing were automated. One or two variable data sets ($n < 100$) could be entered in the calculator then:

descriptive statistics, such as the mean and standard deviation, calculated; graphs such as histogram and scatterplot, displayed; and some regression equations, such as linear, exponential, ln, log, power and inverse, could be calculated. In our study, a graphing calculator with the above capabilities was furnished to Class B.

Class A received a graphing calculator with inferential statistics capabilities. At the time of writing, Casio, Sharp and Texas Instruments offer these calculators for less than $100 in the USA. These calculators have many sophisticated statistical capabilities that include inferential statistics such as hypothesis testing, confidence interval calculations, and one-way analysis of variance.

5. WHY USE THE GRAPHING CALCULATOR?

There are several reasons to use the advanced graphing calculator in an introductory statistics course. The major reasons are access and economics (Kemp, Kissane, & Bradley, 1998). For an investment of less than $100, a student has access to technology at home and in the classroom at all times. A student can do homework and take examinations using the calculator, which is difficult to do with a computer. Many students already own a calculator, especially if they are recent high school graduates. The Advanced Placement examinations in Calculus and Statistics require the use of a graphing calculator. The high school mathematics curriculum has included the incorporation of calculator technology for several years. Of course, the calculator does not have all the capabilities of computer software packages. Data set size and analyses are limited and printing results is not as easy. However, the calculator is a useful tool and teaching aid for introductory statistics.

6. BASIC METHODOLOGY OF THE STUDY

Class A consisted of 22 individuals who completed one section of an introductory statistics course and were provided with a calculator capable of inferential statistics. Class B was 47 individuals who completed another section of the same introductory statistics course with the same instructor. These students were also provided with a calculator, but without the facility for direct inferential statistics. Students enrolled in either Class A or Class B on their own, with no knowledge of the existence of the study. Students in both sections of the course were told that they were being asked to use the calculator in an effort to assess its effectiveness and all students used the calculators on the analyzed examinations. The instructor used the non-inferential calculator overhead while lecturing to both sections. Students were expected to show all the traditional calculations on the inferential examination. The teacher demonstrated the inferential capabilities of the calculator in Class A and did not discuss the inferential capabilities in Class B. Both classes had instruction on other similar statistical capabilities of the two calculator models. The students from each class did not meet or work together to discuss instruction. During examinations, the teacher made sure that the correct calculators were being used. She did not allow the inferential calculator to be used in Class B. The final examination was 2 hours and 45 minutes long and all the problems were compulsory. Both class sections had the same examination though there were four different forms. (See the appendix for an example question with answers.) Every effort was made to keep the experience and evaluation of the two groups as similar as possible.

As an example of the difference in capabilities of the two calculators, consider the exercise of constructing a confidence interval for a proportion, p. Students using the inferential calculator would simply input the confidence level, number of observed successes, and total sample size. Then, the calculator reports an estimate and confidence interval for p. Students using non-inferential calculators would need to first calculate the sample proportion, estimate the standard error, find the appropriate z-value to create the margin of error, and then find the two endpoints of the confidence interval.

For a hypothesis test on a proportion, p, students using the inferential calculator would input the null hypothesis, number of observed successes, total sample size, and type of tail for the test. Then, the calculator reports the test statistic and the p-value for the test. Students using the non-inferential calculators would need to first calculate the sample proportion, estimate the standard error, create the test statistic, and then look up the critical value and p-value for the test. One sample inference problem from the final examination with answers is given in the appendix.

7. DESCRIPTION OF THE DATA

The following variables were collected from the 69 students who completed one of two sections of an introductory statistics class during one spring semester at the university:

- Gender, ethnicity, major;
- Pretest scores (test scores on the first three examinations);
- Inferential test score (test score on the fourth examination);
- Daily grade (a score based on 20 daily quizzes and homework converted to a 100-point scale);
- Final examination score (and answers to individual final examination questions).

All individual examination scores and the daily grade were recorded on a 100-point scale.

Both classes occurred on a Monday/Wednesday afternoon schedule (Class A at 3:30pm and Class B at 2:00pm) with the same instructor and the same examinations. Unfortunately, many variables (both observed and unobserved) confound the study and we cannot separate the "classroom effect" from the "calculator effect." In fact, we observed several differences between Class A and Class B besides the assignment of different calculators for the study. For example, Class A was smaller (22 students) than Class B (47 students) and Class A suffered a greater withdrawal rate of 8 students compared to just 1 withdrawal from Class B. Also, those students remaining in Class A for the entire semester scored significantly higher on their examinations both before and after receiving the calculators for the study.

Nonetheless, there are interesting facets in the data. For example, although students in Class A tended to score higher on both examination 3 (prior to receiving the inferential calculator) and examination 4 (after getting the calculator), they did not show any greater "improvement" in performance after receiving the calculator. We demonstrate these phenomena in the analysis that follows.

Note that the 63 students included in the data analyses (21 from Class A and 42 from Class B) represent only those students who completed the final examination. In addition, sample sizes for the other examinations vary slightly since students were allowed to drop one test score (not the final examination) and indeed three students did not complete examination 4 (one student from Class A and two from Class B). These students provided no information about the effect of an inferential calculator on their examination 4 performance.

8. DATA ANALYSIS

Table 1 gives some summary information on the two groups of students in this study. Recall that students in Class A received an inferential calculator after examination 3 and students in Class B received a non-inferential statistical calculator as described in Section 6. The t-test p-values are for the two-sided pooled-variance t-test of the difference in mean scores. Bartlett's test of the equality of variances (not shown) indicates that the variance of scores are not significantly different for the two classes for the various examinations (except examination 4). The t-test for examination 4 uses Satterthwaite's approximate degrees of freedom. Dotplots (not shown) of the test scores show no significant departure from an assumption of normality.

Students using the inferential calculator scored an average of 10.1 points higher (on a 100-point scale) on the inferential statistics examination (examination 4) than those students using the non-inferential calculator. The two-sided p-value for an unequal-variance t-test of the difference in mean examination 4 scores is 0.03. However, since we also have information on each student's general testtaking ability from the previous examination (examination 3), we can examine the effect of the calculator on examination scores while controlling for a student's examination-taking ability. The plot of the data in Figure 1 illustrates the relationships. In general, students scoring higher on examination 3 also score higher on examination 4.

Table 1. Data Summary

Description	Class A	Class B	
Number of Students	21	42	
Female Students	12 (57%)	20 (48%)	
3rd/4th-year Students	12 (57%)	18 (43%)	
Hispanic Surname	13 (62%)	22 (52%)	
	Avg (sd)	Avg (sd)	t-test p-value
Examination 3	80.0 (16.3)	71.1 (14.3)	0.03
Examination 4	82.3 (13.7)	72.2 (21.3)	0.03
Final Examination	77.2 (10.5)	74.8 (10.3)	0.38
Daily Grade	85.0 (17.7)	85.2 (18.4)	0.97

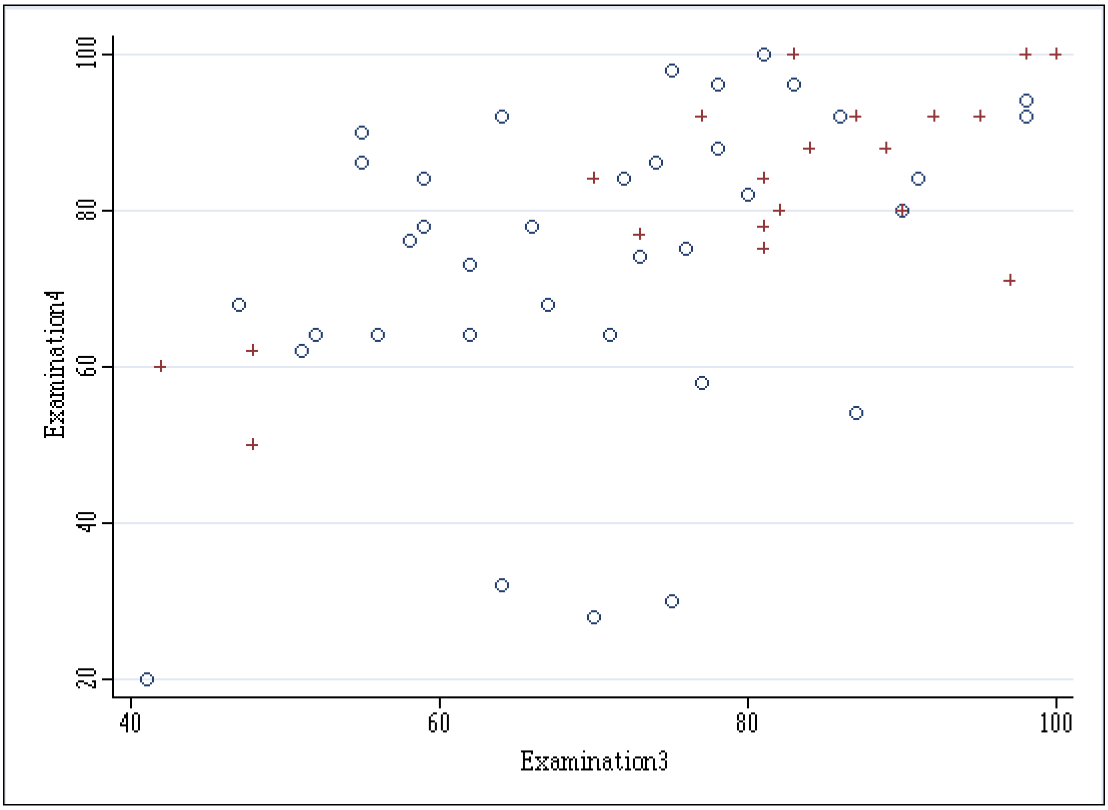

Figure 1: Examination 4 scores by Examination 3 scores. Class A (students using the inferential calculator) marked as "+"

The following linear regression model was fitted to the data:

Examination4 = Intercept + E*Examination3 + C*Calculator.

Here,

Examination4 = score on Examination 4 (inferential topics),

Calculator = 1 for Class A (inferential calculator) and 0 for Class B,

Examination3 = score on Examination 3 (test score prior to obtaining calculators for the study).

The regression results are recorded in Table 2.

Table 2. Least squares regression results to estimate calculator effect while adjusting for prior examination scores

Dependent Variable: examination4

Coefficient	Parameter Estimate	Standard Error	T for Ho: Parameter=0	P-Value: Prob > |T|
Intercept	33.640	12.210	2.76	0.008
E (Examination3)	0.608	0.146	4.17	0.000
C (Calculator)	2.690	4.716	0.57	0.571

R-Square = 28.34%

The examination 3 scores are, as expected, a strongly significant predictor of the average examination 4 score (p-value < 0.0001). However, once the model is adjusted for this measure of examination performance (examination 3 score), the type of calculator used is not at all statistically (or practically) significantly related to performance on the inferential statistics examination (examination 4). A similar analysis with the final examination score as the response variable indicates no significant difference between the two groups of students either conditionally (adjusting for previous scores on examination 3) or unconditionally (as already seen in Table 1 where the p-value for a t-test of the difference in average final examination scores was 0.38).

We should point out that, generally, there is a lot of variability in these data and even the examination 3 scores only account for about 28% of the variability in examination 4 scores when using the linear regression model fit in Table 2. The relationship between examination 3 and examination 4 scores appears fairly linear for both groups, but there is a lot of scatter in the data. A plot of the residuals does not indicate the presence of a non-linear pattern.

To look further for any effect of the inferential calculator, we took a deeper look at students' performance on specific examination questions. See, for example, the sample final examination question in the Appendix. In this question, Parts I, II, and V could be considered conceptual and III, IV, and VI calculations (with some overlap). We found that both classes had trouble with the "calculations" while at the same time answering "concept" questions fairly well.

9. DISCUSSION AND CONCLUSIONS

In general, we observed no difference between the two groups of students in their performance on examinations on inferential topics in an introductory statistics course. In particular, use of an inferential calculator that performs many of the intermediate steps for calculating confidence intervals and p-values does not appear to be related to student performance. The study size was small and the design did not allow for the separation of the "calculator effect" from the "classroom effect" (a confounding factor). However, it is interesting to note that although Class A generally performed better on examinations prior to receiving inferential calculators for the study, these same students did not perform significantly or practically better on inference topics after using the calculator. The inferential calculator did not appear to give students any clear advantage (or disadvantage) in their performance on examinations. This study suggests that use of the inferential calculator needs to be explored further as a benefit for student-learning in an introductory statistics courses. We encourage others who may teach within an infrastructure that could allow a randomized experiment to complete a study and share their results.

Of course, if the inferential calculator is required in an introductory statistics course, the instructor would be able to spend much less time on computation and more time on gathering and analyzing real-world data. Indeed, as classroom computer use has changed how statistics has been taught during the last 20 years (as computers have changed the practice of statistics), inferential calculators can also change how introductory statistics is taught.

ACKNOWLEDGEMENTS

This research was completed at the University of Texas at San Antonio.

REFERENCES

Campbell, P., & Stewart, E.L. (1993). Calculators and computers. In R. Jensen (Ed.), *Early Childhood Mathematics, NCTM Research Interpretation Project* (pp. 251-268). New York: Macmillan Publishing Company.

Carlson, M. P. (1995). A successful transition to a calculator integrated college algebra curriculum: Clues, surveys, and trends. In P. Bogack (Managing Ed.), E. D. Fife, & L. Husch (Eds.), *Proceedings of the Seventh Annual International Conference on Technology in Collegiate Mathematics* (pp. 73-77). Reading, MA: Addison-Wesley Publishing Group.

Cobb, G. W. (1993). Reconsidering statistics education: A National Science Foundation conference, *Journal of Statistics Education, 1*(1). [Online: www.amstat.org/publications/jse]

Demana, F., & Waits, B. K. (1990). Implementing the Standards: The role of technology in teaching mathematics, *Mathematics Teacher*, *83*, 27-31.

Dunham, P. H. (1993). Does using calculators work? the jury is almost in. *UME Trends*, *5*(2), 8-9.

Dunham, P. H. (1996). Looking for trends: What's new in graphing calculator research? In P. Bogack (Managing Ed.), E. D. Fife, & L. Husch (Eds.), *Proceedings of the Eighth Annual International Conference on Technology in Collegiate Mathematics* (pp. 120-124). Reading, MA: Addison Wesley Publishing Company.

Dunham, P. H., & Dick, T. P. (1994). Research on graphing calculators. *Mathematics Teacher*, *87*, 440-445.

Fey, J. T., & Good, R. A. (1985). Rethinking the sequence and priorities of high school mathematics curricula. In C. R. Hirsch & J. Zweng (Eds.), *The secondary school mathematics curriculum, 1985 yearbook* (pp. 43-2). Reston, VA: National Council of Teachers of Mathematics.

Francis, B. (1997). Graphs of binomial and Poisson distributions on a graphical calculator. *Teaching Statistics*, *19*(1), 24-25.

Garfield, J. (1995). How students learn statistics. *International Statistical Review*, *63*(1), 25-34.

Gilchrist, W. (1986). Teaching statistics to the rest of humanity. In R. Davidson & J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics* (pp. 494-497). Victoria, British Columbia, Canada: University of Victoria Conference Services.

Gnanadesikan, M., Scheaffer, R. L., Watkins, A. E., & Witmer, J. R. (1997). An activity-based statistics course. *Journal of Statistics Education, 5*(2).

Graham, A. (1996a). Following form with a graphic calculator. *Teaching Statistics*, *18*(2), 52-55.

Graham, A. (1996b). The TI-83 and TI-92 calculators. *Teaching Statistics*, *18*(3), 94-95.

Graham, A. T., & Thomas, M. O. J. (2000). Building a versatile understanding of algebraic variables with a graphic calculator. *Educational Studies in Mathematics*, *41*(3), 265-282.

Harvey, J. G., Waits, B. K., & Demana, F. D. (1995). The influence of technology on the teaching and learning of algebra. *Journal of Mathematical Behavior*, *14*, 75-109.

Hembree, R., & Dessart, D. J. (1986). Effects of hand-held calculators in precollege mathematics education: A meta-analysis. *Journal for Research in Mathematics Education*, *17*, 83-89.

Hennesey, S., Fung, P., & Scanlon, E. (2001). The role of the graphic calculator in mediating graphing activity. *Education in Science and Technology*, *32*, 267-290.

Hogg, R.V. (1992). Towards lean and lively courses in statistics. In F. Gordon & S. Gordon (Eds.), *Statistics in the twenty-first century [MAA Notes No. 26]* (pp. 3-13). Washington, DC: Mathematical Association of America.

Kemp, M., Kissane, B., & Bradley, J. (1998). Learning undergraduate statistics: The role of the graphics calculator. In B. Parker (Ed.), *Proceedings of the International Conference on the Teaching of Mathematics* (pp. 176-178). Reading, MA: John Wiley and Sons, Inc.

Mittag, K. C., & Eltinge, E. (1998). Topics coverage in statistics courses: A national Delphi study. *Research in the Schools*, *5*(1), 53-60.

Moore, D. S. (1992). Teaching statistics as a respectable subject. In F. Gordon, & S. Gordon (Eds.), *Statistics in the Twenty-First Century [MAA Notes, No. 26]* (pp. 14-25). Washington, DC: Mathematical Association of America.

Moore, D. S., Cobb, G. W., Garfield, J., & Meeker, W. (1995). Statistics education fin de siecle. *The American Statistician*, *49*(3), 250-260.

Neter, J. (1989). Undergraduate statistics service courses in the years ahead. In J. Bolinger & R. Brookmejer (Eds.), *American Statistical Association, 1989 Proceedings of the Section on Statistical Education* (pp. 29-31). Alexandria, VA: American Statistical Association.

Pirie, W. R. (1989). Undergraduate statistics programs: Challenges for the 1990's. In J. Bolinger & R. Brookmejer (Eds.), *American Statistical Association 1989 Proceedings of the Section on Statistical Education* (pp. 32-33). Alexandria, VA: American Statistical Association.

Quesada, A. R., & Maxwell, M. E. (1994). The effects of using graphing calculators to enhance college students' performance in precalculus. *Educational Studies in Mathematics*, *27*, 205-215.

Rinaman, W. C. (1998). Revising a basic statistics course, *Journal of Statistics Education*, *6*(2). [Online: www.amstat.org/publications/jse]

Singer, J. D., & Willet, J. B. (1990). Improving the teaching of applied statistics: Putting the data back into data analysis. *The American Statistician*, *44*(3), 223-230.

LINDA BRANT COLLINS
Department of Statistics
University of Chicago
5734 S. University Ave
Chicago, IL 60637
United States of America

APPENDIX: FINAL EXAMINATION INFERENCE QUESTION WITH ANSWERS

A statistics professor surveys 60 students and finds that 12 are left-handed. Use a 0.05 level of significance to test the claim that these students come from a population in which the percentage of left-handed people is greater than 10%.

Part I: What is the correct null and alternative?

a) H0: $p \leq 0.10$ H1: $p > 0.10$

b) H0: $p > 0.10$ H1: $p \leq 0.10$

c) H0: $p = 0.10$ H1: $p \neq 0.10$

d) H0: $p \neq 0.10$ H1: $p = 0.10$

e) none of these

Answer: a) H0: $p \leq 0.10$ H1: $p > 0.10$

Part II: Which of the following is true?

a) This is a two-tailed test.

b) This is a left-tailed test.

c) This is a right-tailed test.

Answer: c) This is a right-tailed test.

Part III: What is the test statistic?

a) -1.94

b) -2.58

c) 2.58

d) 1.94

e) none of these

Answer: c) 2.58

Part IV: What is the critical value?

a) 1.645

b) 2.575

c) 1.96

d) none of these

Answer: a) 1.645

Part V: What is the conclusion?

a) There is not sufficient evidence to reject the claim that the proportion is more than 0.10.

b) There is not sufficient evidence to support the claim that the proportion is more than 0.10.

c) There is sufficient evidence to reject the claim that the proportion is more than 0.10.

d) There is sufficient evidence to support the claim that the proportion is more than 0.10.

e) none of these

Answer: d) There is sufficient evidence to support the claim that the proportion is more than 0.10.

Part VI: What is the p-value for this problem?

a) 0.0049

b) 0.4951

c) 0.4738

d) 0.0262

e) none of these

Answer: a) 0.0049