Teaching Statistics from the Operating Table: Minimally Invasive and Maximally Educational

Amy S. Nowacki
Cleveland Clinic and Cleveland Clinic Lerner College of Medicine

Abstract

Statistics courses that focus on data analysis in isolation, discounting the scientific inquiry process, may not motivate students to learn the subject. By involving students in other steps of the inquiry process, such as generating hypotheses and data, students may become more interested and vested in the analysis step. Additionally, such an approach might better prepare students to tackle real research questions outside of the statistics classroom. Presented here is a classroom activity utilizing the popular Hasbro board game Operation, which requires student involvement in the entire research process. Highlighted are ways this activity uncovers a number of research issues. A number of categorical and continuous variables are collected, making the activity amenable to a variety of statistical investigations and thus easy to imbed into any curriculum. Designed to mimic a real-world research scenario, this fun activity provides a guided yet flexible research experience from start to finish.

1. Introduction

A theme emerged from meetings with former students during their research year, which can be summarized as follows: I believe I know what to do once I have the dataset, but I am not sure how to get the dataset. You always gave us our datasets.

Inadvertently, the format of my course reinforced the mistaken equivalence of statistics with data analysis and failed to convey the larger investigative cycle. If students are always supplied with clean formatted datasets upon which to apply the analytical technique(s) and occasional brief

descriptions of where the data originated, without enough detail imparted that the student could generate his or her own dataset, could such a course format actually impede students from understanding the research process?

The idea that statistical problem solving is an investigative cycle has been described in detail by Wild and Pfannkuch (1999). Based on the Problem, Plan, Data, Analysis, Conclusions (PPDAC) model (MacKay and Oldford 1994), the cycle consists of five phases. First, grasping system dynamics (context knowledge) and defining the problem. This is followed by thoughtful planning of the measurement system (what to measure and how to measure it) and sampling design (what units to take measurements on). Next is the data collection, management and cleaning. Then, data exploration, planned analyses, unplanned analyses and hypothesis generation occurs. Finally, interpretation, conclusions, new ideas and communication of results takes place. The conclusions from the investigation feed into an expanded context-knowledge base which can then inform any actions.

Viewing statistical problem solving as an investigative process was again reiterated in both American Statistical Association (2005, 2012) GAISE Reports. In the College Report (2012), a recommendation is made that activities should include design, data collection and analysis so students can see the whole process at work. Similarly, in the PreK-12 Report (2005), the investigative process is described as a framework with four components: formulate questions, collect data, analyze data and interpret results. Students should understand the parts of the process through which statistics works. First, one must clarify the problem at hand and ensure that the question(s) can be answered with data. An important next consideration is: how to obtain or generate data? In order to teach statistical thinking, it is essential to emphasize the importance of data production (American Statistical Association 2012). Students should be taught to appreciate that it is difficult and time-consuming to formulate problems and to get quality data to address the right questions. Most students may not realize this until they go through this experience themselves. Optimally, this encounter should be in a supportive learning environment (such as the classroom) and not on the job where high stakes are often encountered.

Thus, I sought to develop a new activity for the first session of an introductory statistics course focusing on the investigative cycle. The goal of this session was to prepare students for study start-up. The specific learning objectives were to: (1) use the literature to develop research questions; (2) identify components of a clearly defined research question; (3) compare and contrast types of studies (experimental vs. observational); (4) define independent and dependent variables; (5) classify types of variables (numeric, categorical, binary, etc.); and (6) build a database. This article aims to describe a classroom activity which utilizes the popular Hasbro board game Operation to illustrate statistical problem solving as an investigative process.

Appealing to my primary student audience of medical students, the board game Operation was selected to serve the purpose of data generation. Operation is a battery-operated game of physical skill that tests players' hand-eye coordination and fine motor skills. It consists of an "operating table," lithographed with a comic likeness of a patient (nicknamed "Cavity Sam") with a large red light bulb for his nose. In the surface are a number of openings, which reveal cavities filled with fictional and humorously named ailments made of plastic. The general gameplay requires players to remove these plastic ailments with a pair of tweezers. However, if the tweezers touch

the metal edge of the opening during the attempt, a buzzer sounds, Sam's nose lights up red, and the player loses the turn.

2. Methods

2.1 Setting

This activity was initially developed for a required introductory biostatistics course comprised typically of 32 second-year medical students in a program designed to train physician investigators. The class met for 2 hours twice a week over a period of 8 weeks. The classroom was equipped with a computer (laptop option) projected onto a screen in the classroom. There were 8 large round tables (1 table for each group of 4 students). Students were encouraged to bring laptops for real-time data collection. The activity was designed to be implemented on the first day of class (with appropriate prior reading). This activity has been implemented during the past 3 offerings of the course.

2.2 Student Preparation

Students are asked to prepare for class by completing three brief reading assignments. The first provides an introduction to the following statistical topics: flow of a research study, population vs. sample, observational vs. experimental study designs, randomization, independent vs. dependent variables, covariates, and qualitative vs. quantitative variables. Such content can typically be found in the first chapter of any statistical textbook. Specifically, I utilize section 1.1 Introduction to the Practice of Statistics from the Sullivan textbook (2004). The second two are journal articles addressing the controversial topic of what characteristics make for a successful surgeon. One article reviews the available data on the impact of left-handedness on surgical training (Tchantchaleishvili and Myers 2010). The other article aims to identify factors (hand dominance, gender and experience with computer games) that influence surgeons' performance, as measured by a virtual reality computer simulator for laparoscopic surgery (Grantcharov, Bardram, Funch-Jensen, and Rosenberg 2003).

3. Activity Description

To begin the activity, arrange students into groups and do not have board games visible. Ask each group to discuss the assigned readings regarding what characteristics make a successful surgeon and have each group generate at least two hypotheses that they would like to test motivated by the articles. This provides an opportunity to discuss the features of good testable research hypotheses: specification of the subject group, the treatment this group will receive (exposure of interest), the outcome measure, and the comparison or control group. Have the class come together for a large group discussion where hypotheses are shared. This step is important as it emphasizes that the students generated the research question(s) (GAISE recommendation). Next, inform the students that based on their suggestions, they will be designing and conducting a study to investigate the impact of both handedness and gender on surgeon success rates. While it is possible students will not have identified these variables, it is likely, since the articles assigned as reading prime the students about comparing left and right handed surgeons as well as male and female surgeons.

Use the handout in Appendix A to frame the following discussion. Together as a class, design the study, both defining terms and identifying the population (surgeons) and sample (surgeons in training (class name) on (class date)). Review common study designs and their features (either specific types such as cohort, case-control, clinical trial or keep it general such as experimental versus observational) and discuss the designs that provide the highest level of evidence and why. This allows a brief introduction to the consideration of randomization and ethics, blinding, and independent (one surgery per surgeon) versus dependent (multiple surgeries per surgeon) observations. Because it is unethical to randomize which hand a surgeon uses in performing surgery (think of the poor patients!) and not feasible to randomize a surgeon to gender, guide the class toward a prospective cohort study or more generally an observational study.

Next, define terms and identify the independent and dependent variables. Handedness and gender are the primary independent variables, but solicit other covariates of interest that may be collected (e.g., type of surgery, experience level of surgeon, etc.). Emphasize the importance of clearly defining the outcome variable and discuss various definitions of a successful surgery. Ultimately declare that if during the surgery the surgeon avoids touching the edge of the opening with the surgical tool, it will be recorded as a success, but if a buzzer sounds and the patient's nose lights up red, a failure. Students typically make the connection to the game Operation at this point and laugh thinking it is a joke. Secondary outcomes to be collected are length of surgery and success on first attempt. Have each group write down a definition of length of surgery and share with the class to emphasize how seemingly obvious variables are subject to interpretation and hence the importance of a protocol. Define length of surgery as the time from when the surgeon first touches the tweezers until replacement on the operating table. Finally, to conclude the study design discussion, ask students to classify each variable to be collected as categorical or quantitative. Be sure to emphasize why this step is important as many students do not yet realize that the types of variables involved will determine the analytical approach.

Optionally, this activity can be used to motivate an introduction to statistical software. Our institution utilizes JMP statistical software and this activity serves as a framework for discussion of database structure and how to record the data and organize the data for analyses (row = observation, column = variable). Students follow along on their laptops as we build the database step-by-step; alternatively, this could be the basis of a tutorial or assignment. Having already had completed the variable classification exercise (Appendix A), students now use the information as they must define each variable within JMP as either numeric or character (aka categorical). Additionally demonstrated (but optional) are the following skills: importing an EXCEL file, assigning labels, data range checks, subsetting and creation of new variables. For illustration purposes, a new variable called QUICK is created by dichotomizing the TIME variable at 5 seconds using the formula editor. Saving the data file is demonstrated and performed by each student so that they all have identical copies of the empty database, a critical element to ease data collection and collation.

At this point, the student-surgeons are ready to perform surgery and collect data. Reveal the Operation game boards. Begin with a few instructions (recommended they are projected on a large screen during review) as outlined in Figure 1. In an attempt to standardize the activity, Figure 2 provides a study protocol (recommended it be projected on a large screen and a hard copy inside each game box). Review the study protocol together in detail. Note that in step 3, the

ailment to remove for each surgery must be determined. The original version of the game has 12 ailments. Brain Freeze was added in 2004, when Milton Bradley allowed fans a chance to vote on a new piece to be added to the original game. Thus original versions after 2004 have 13 ailments, as do all modern versions. It is recommended that one use a random number generator to obtain an integer between 1 and 12 or 13 to choose the ailment. Many convenient websites are available for computers or free apps for smartphones (examples provided in instructions, Figure 1). The random number can then be matched to the surgical ailment on provided code sheets (Figure 3) (recommended to insert a hard copy inside each game box). From experience, it is suggested that you remind the students of the variable definitions (particularly the outcomes: TIME, ATTEMPT and OUTCOME) and refer them to the previous handout with variable definitions (Appendix A). Finally, begin the data collection.

Figure 1. Instructions for Students on Activity Set-up

<u>Instructions</u>

1. Select one group members' computer for data collection

2. Select one group members' computer/smartphone to use as a random number generator to determine surgical ailment
   - computer: http://www.randomnumbergenerator.com/
   - Smartphone: Random Number Generator + App (free)

3. Select one group members' computer/smartphone to use as a timer
   - computer: http://www.online-stopwatch.com/full-screen-stopwatch/
   - Smartphone: stopwatch feature of clock

Figure 2. Study Protocol

<u>Study Protocol</u>

1. Identify the operating **surgeon** (scrub-in)

2. The person on the surgeons' right is the **data collector** who records baseline data (SURGEON, HAND, GENDER and EXPERIENCE)

3. The person on the surgeons' left is the **study coordinator** who identifies surgery ailment (use ailment code sheet)

4. All other group members observe the surgery

5. Data collector records AILMENT variable

6. Study coordinator prepares to time the surgery (seconds), determine success of first attempt (yes/no) and declare surgery outcome (success/failure)

7. Surgeon performs operation on assigned ailment

8. Data collector records TIME, ATTEMPT and OUTCOME variables

9. Study coordinator becomes new operating surgeon

10. Repeat steps 2 – 9

11. When instructed, save JMP data file and email to instructor

Figure 3. Code Sheet for Surgical Ailment Assignment

A   Original Version

⌁ Parts Chart ⌁

Shape	Name	Number
	Adam's Apple	1
	Broken Heart	2
	Wrenched Anle	3
	Butterflies in the Stomach	4
	Spare Ribs	5
	Water on the Knee	6
	Funny Bone	7
	Charley Horse	8
	Writer's Cramp	9
	Ankle Bone Connected to the Knee Bone	10
	Wish Bone	11
	Bread Basket	12
	Brain Freeze	13

B   Silly Skill Game Version

⌁ Parts Chart ⌁

Shape	Name	Sound	Number
	Bird Brained	Chirping Bird	1
	Runny Nose	Sneeze	2
	Ringing in the Ear	Bell Ringing	3
	Frog in the Throat	Frog Ribbit	4
	The Giggles	Laughing	5
	Burp Bubbles	Bubbles/Burp	6
	Bad Plumbing	Toilet Flush	7
	Time on his Hands	Clock Ticking	8
	Phone Finger	Cell Phone Dialing	9
	Toxic Gas	Passing Gas	10
	Water on the Knee	Dripping Water	11
	Pulled Muscle	Stretching/Boing!	12
	Dog Tired	Barking Dog	13

In addition to the lessons learned in actually completing the activity, the resulting dataset contains both continuous and categorical variables and thus can be utilized for a number of other sessions. Thus far, the dataset has served as a basis for exploration of descriptive statistics (measure of center and variability, impact of outliers, building a *Table 1* for a journal article), estimation of a mean or proportion and confidence intervals, one-sample test of a mean or proportion, two-sample t-test, one-way ANOVA, multiple testing issues, chi-square and Fisher's exact tests (see Appendix B for examples). The dataset is not the focus of those sessions, but it serves as a nice example easily incorporated into in-class and/or at-home exercises or exam questions. Be sure to revisit the dataset minimally to answer the primary research question (a chi-square test to compare the success rates and a t-test to compare the length of surgeries among the genders).

The astute reader might find it unusual that the analytic plan does not include the comparison of success rates among left and right handed surgeons. This comparison is actually a red-herring (intended to be misleading) and reveals study design issues. The problem with hand dominance is that only approximately 10% of the population is left-handed; thus, depending on the class

size, there are typically few or no left-handed students. This realization may occur during the actual study design discussion (though rarely since it is the first day of class), the data collection process (when a group has only right-handers for example) or during data analyses (my experience). This promotes discussion of concepts such as risks of observational studies, rare exposures and cohort studies, oversampling techniques, etc. A couple of options are available: (1) decide to drop this research question; (2) oversample left-handers, (3) simulate data (parametric or bootstrap); or (4) declare a success rate for left-handed surgeons and perform a one-sample test of a proportion to see if right-handed surgeons differ. While the students often feel disappointed that they cannot address this research question as they had initially planned, this component is kept in the activity as an important lesson demonstrating the limitations of studies. Researchers do their best to address study limitations, but minimally they must at least be acknowledged.

Described below are two potential add-ons to this activity, each with the goal of introducing a specific object lesson. The first addresses the statistical idea of confounding.. When forming groups of students, select at least two groups to be of single sex (i.e., all males or all females). Distribute to the groups of students different versions of the board game representing various difficulty levels. The original version is for ages 6+ whereas some newer versions are advertised as having "Large Openings with Easy-Grab Pieces" for ages 4+. The age recommendations are always a good hint at determining the difficulty level. The issue of confounding can be introduced when you select the single sex groups as representative of the class and then compare surgery success rates between the genders. If, for example, the females had the easier version, they typically have a higher success rate. Declare females the better surgeons. Then reveal that each version has a different difficulty level and have each group check their version. In my experience, the males adamantly claim injustice when they discover that they had a more difficult version. Confounding point made. The primary research questions, however, can still be addressed by using the data from the other mixed gender groups for which the difficulty level should impact both genders equally. When using different versions of the game, comparisons among the various ailments is no longer possible (each version has its own unique ailment list). There is, however, one option which preserves the ailment comparison while allowing different skill levels. In August 2008, Hasbro released a "Silly Skill Game" version which features two difficulty levels. Level 1 is for younger or inexperienced players where they are less likely to get buzzed if the tweezers lightly touch the metal side of a cavity and level 2 for older or more experienced players where they get buzzed right away if one touches the metal side of a cavity. While tempting, I do not recommend this approach because these two settings require the use of the Funatomy sound effects (more discussion on this later).

The second add-on is a data cleaning lesson. Whether the class creates a database together as previously described or the instructor distributes an empty database to each student group, the variables will be standardized and allow for data collation. What will not be standardized is how students enter the actual raw data values. For example, the hand of the surgeon may be entered as left, Left, L or 1 by the various groups. Thus, when collated and redistributed to the students for other sessions (recommended as an add-on to a descriptive statistics session), students quickly realize the issue. Of course, the instructor can make the data messier intentionally to emphasize any particular point in data management. In my experience, this has been unnecessary as the data tend to be messy enough when submitted by the students.

4. Discussion

4.1 Implementation Issues

The time devoted to the data collection portion of the activity depends on the size of the class, the number of groups, and the number of board games. The number of board games required depends on both the class size and the amount of time available to devote to data collection. The first year this activity was piloted, I had only one board game for the entire class of 32 students and 50 minutes to collect data. This resulted in a small sample size (n = 64 surgeries) as time only permitted each student to perform surgery twice. This result is optimistic as I recorded the data myself which streamlined the process but did not allow the students to experience this important step. After the first implementation of the activity, I increased the number of board games to four and allocated data recording responsibilities to the students. A rough estimate of the average time per surgery is 2 minutes. Thus if you have four game boards, one can expect to have approximately 60 surgeries completed in 30 minutes. Minimally, I suggest allowing enough time for each student to perform surgery at least once to maintain engagement and, if possible, for the class as a whole to perform at least 100 surgeries. While no formal sample size calculation is being performed (this is the first day of class after all), this recommendation allows one to collect enough data so as not to be completely underpowered.

Cost is a consideration as board games range from $5 to $40 depending on the version of the game. Costs can be managed by slowly building inventory or the frugal shopper can utilize eBay and Christmas sales where substantial savings are often encountered. Also, remember that batteries must be purchased for each game unit.

One must also consider the noise factor: as the number of board games increases, so too does the number of groups talking and buzzers indicating failed surgeries. The "Silly Skill Game" version features 13 different Funatomy sound effects for each of the different ailments. This version has two difficulty settings, as previously mentioned, where the game automatically generates the noise associated with a particular ailment in random order, thus removing the need for random number generation. While the sounds are comical, I recommend not using these options as the classroom noise level becomes disruptive. Rather, with this version, I use a third setting option which turns off the Funatomy sound effects and allows the removal of any Funatomy part quietly.

Lastly, the board games need to be stored when not in use. Have a storage plan and consider the transportation of board games to and from the classroom as a single person can only carry approximately 5-6 games comfortably.

4.2 GAISE Recommendations

The described activity adheres to many of the GAISE desirable characteristics of class activities (American Statistical Association 2012) such as grounding activities in the context of real problems (mimic real-world scenarios) and ensuring that questions used with data sets are of interest to students. While there is no illusion that playing the board game Operation mimics real surgery, the focus of realism is on the research process: utilizing the literature to develop a

research question, designing a study and creating a database is exactly the same process employed in traditional research. By grounding the activity in the context of a real problem, the data collection is for the purpose of answering a question, not simply to practice collecting data. As the assigned preparatory readings are from the current medical literature, students appreciate the realistic context of the investigation. Relevance is of utmost importance, particularly when training primarily consumers of statistics. As many of our medical students are considering a career in surgery, exploring characteristics of a successful surgeon has been of great interest and more relevant to medical students than card games, dice rolling, puzzles or prepackaged textbook datasets. While specifically designed for medical students, students of other disciplines (premed, biomedical sciences, etc.) would also benefit from such an activity as the study start-up process is universal.

Another GAISE suggestion is to not use activities that lead students step by step through a list of procedures, but allow students to discuss and think about the data and the problem. Thus, this activity is student driven, but with direction. Students experience the whole research process, with emphasis on the interconnectedness among the research question, study design, data collection, analyses, and interpretation. It has been reported that when students have input into the research question and/or the design of a study, all other aspects of the research process have fun elements. Those include collecting and analyzing data, but more importantly discerning the meaning of the data and drawing conclusions (Jarrett and Burnley 2010). One could describe this activity as inquiry learning in which students have opportunities to answer their own research questions, collect their own data, collaborate with peers, and draw meaning from the data. Not only seeking student input in the decision process, but also the mere act of incorporating the whole research process is important. Encouragement for incorporating the scientific process of interdisciplinary data analysis as statistics is actually practiced is growing (Cummiskey, Kuiper, and Sturdivant 2012). Developers of other such in-class experiments note that such activities allow the opportunity for students to see for themselves some of the practical issues that can arise when carrying out research studies (Brophy and Hahn 2014). As such, students are better prepared to tackle real research questions outside of the statistics classroom.

GAISE also suggests that well-designed activities should provide students an opportunity to practice collecting and entering raw data. Since students are playing the board game (aka performing surgery) in-class, they are collecting real-time data. Collection and data analysis is at the heart of statistical thinking. Data collection promotes learning by experience and connects the learning process to reality (Snee 1993). Involving students in data collection should stimulate greater interest in the investigation. Such techniques are useful to improve student attendance and engagement, but the impact on statistical reasoning skills is still being debated (Pfaff and Weinberg 2009). As a bonus, the activity serves as a lesson on how to structure data, a topic often inadequately addressed in introductory courses (Cummiskey et al. 2012).

A final GAISE suggestion is to use subsets of variables in different parts of the course, but integrate the same dataset throughout. Described are a number of variables, both categorical and continuous, that can be collected during this activity allowing the dataset to serve multiple purposes as a variety of statistical investigations are applicable. A handful of examples are presented in Appendix B. As noted by others utilizing this technique, later use of the dataset in

the course saves time as it requires little description of the study as students immediately identify the context (Schwartz 2013).

Conclusion

While fun for students, this activity is also enjoyable and stimulating for instructors as one never knows the exact results from a given classroom experiment. Overall, the activity is an engaging and memorable approach for teaching general research skills and statistical reasoning. Unlike a research assignment where the control is solely in the hands of the instructor (they come up with the questions, the study design, data, etc.), this is a research experience where the control is shared and the students contribute in the decision process. Students discover the investigative cycle from generation of hypotheses to data analyses and results preparing them to conduct research outside the statistical classroom.

Appendix A

What Makes a Successful Surgeon?
_____ et al. (2014)
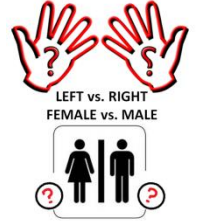
Primary Aim: Investigate whether handedness or gender is associated with being a successful surgeon.

1.   Explain the difference between a population and a sample.

2.   Identify the population and the sample in this study.

3.   Define observational and experimental studies.

4.   Which provides the highest level of evidence?

5.   What is the study outcome?

6.   What is the variable(s) of interest?

7.   What other variables are important to collect?

8.   For each of the variables to be collected in this study, classify by circling all taxonomy that applies.

Variable	Description
SURGEON	surgeon name
HAND	hand performing surgery (left, right)
GENDER	gender of surgeon (male, female)
AILMENT	ailment removed during surgery (broken heart, frog in throat, etc.)
EXPERIENCE	has surgeon ever played the board game Operation (yes, no)
OUTCOME	ailment removed without nose buzzer or light (success, failure)
TIME	length of surgery (time from when surgeon first touches tweezers to replacement on operating table)
ATTEMPT	successful first attempt (yes, no)

- **SURGEON**

    qualitative   quantitative   nominal   numeric   categorical   binary

- **HAND**

    qualitative   quantitative   nominal   numeric   categorical   binary

- **GENDER**

    qualitative   quantitative   nominal   numeric   categorical   binary

- **AILMENT**

    qualitative   quantitative   nominal   numeric   categorical   binary

- **EXPERIENCE**

    qualitative   quantitative   nominal   numeric   categorical   binary

- **OUTCOME**

    qualitative   quantitative   nominal   numeric   categorical   binary

- **TIME**

    qualitative   quantitative   nominal   numeric   categorical   binary

- **ATTEMPT**

    qualitative   quantitative   nominal   numeric   categorical   binary

Appendix B

Examples of Teaching Applications Using the Data from the Activity

(1) Exploration of Descriptive Statistics

It is of interest to know if the success of a surgery is affected by the gender of the surgeon. Therefore, we are interested in comparing two groups: surgeries performed by male surgeons and surgeries performed by female surgeons. Our main outcome of interest is the difference in the proportion of successful surgeries between these two groups. To demonstrate that an observed difference is likely due to the gender of the surgeon, one wants to demonstrate how the two groups are similar/dissimilar with regards to basic demographics and potential confounders. This can be done using descriptive statistics. An empty Table 1 is provided; fill in the table with the appropriate descriptive statistic describing the two groups. Clearly label all entries and be ready to defend your selections. For now, ignore the p-value column.

Table 1. Characteristics of Performed Surgeries

	Male	Female	p-value
n*	19	17	
right-handed surgeon	17 (89%)	15 (88%)	
ailment			
bird brained	2 (11%)	0 (0%)	
runny nose	0 (0%)	3 (18%)	
ringing in the ear	2 (11%)	1 (6%)	
frog in the throat	4 (21%)	0 (0%)	
the giggles	1 (5%)	4 (24%)	
burp bubbles	1 (5%)	0 (0%)	
bad plumbing	1 (5%)	1 (6%)	
time on his hands	1 (5%)	1 (6%)	
phone finger	1 (5%)	1 (6%)	
toxic gas	1 (5%)	1 (6%)	
water on the knee	1 (5%)	2 (12%)	
pulled muscle	1 (5%)	3 (18%)	
dog tired	3 (16%)	0 (0%)	
experienced surgeon	15 (79%)	12 (71%)	
Length of surgery** (seconds)	8.1 (3.0, 42.2)	4.9 (1.0, 17.0)	
successful first attempt	11 (58%)	14 (82%)	

* Data missing for 4 observations (2 male; 2 female)

** median (minimum, maximum)

(2) One –sample Test of a Binomial Proportion:

60% of the surgeries in the original study were successful. Does the proportion of successful surgeries in our study significantly differ?
To test this, the null and alternative hypotheses are set as:

$H_0: \pi = 0.60$ versus $H_1: \pi \neq 0.60$

where π represents the population proportion of successful surgeries.

$$\hat{p} = \frac{29}{36} = 0.8056$$

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.8056 - 0.60}{\sqrt{0.60(1-0.60)/36}} = 2.52$$

- Enter a value in three of the four text boxes.
- Leave the fourth text box blank.
- Click the **Calculate** button to compute a value for the blank text box.

Standard score (z) 2.52
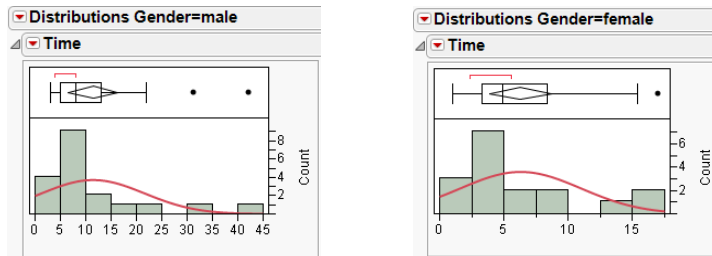Cumulative probability: P(Z ≤ 2.52) 0.99413
Mean 0
Standard deviation 1

P = 2*(1-0.9941) = 2*0.0059 = 0.0118
Reject the null and conclude that the proportion of successful surgeries is higher in our study (81% vs. 60%, p = 0.01).
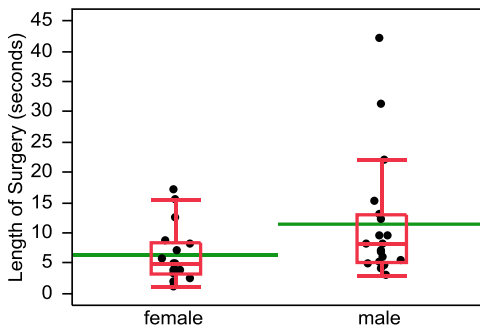
(3) Wilcoxon rank-sum test

Does the length of surgery differ among male and female surgeons?

To test this, the null and alternative hypotheses are set as:

$H_0: M_{male} = M_{female}$ versus $H_1: M_{male} \neq M_{female}$

Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
female	17	247.5	314.5	14.6	-2.1
male	19	418.5	351.5	22.0	2.1

2-Sample Test, Normal Approximation

S	Z	Prob>|Z|
247.5	-2.11	0.0350*

Reject the null and conclude that the length of surgery differs among the male and female surgeons (p = 0.04); with male surgeons having a longer length of surgery (median 8.0 vs. 4.9 seconds respectively).

References

American Statistical Association (2005), "Guidelines for Assessment and Instruction in Statistics Education: PreK-12 Report." Retreived November 24, 2014 from http://www.amstat.org/education/gaise/

American Statistical Association (2012), "Guidelines for Assessment and Instruction in Statistics Education: College Report." Retreived November 24, 2014 from http://www.amstat.org/education/gaise/

Brophy, C., and Hahn, L. (2014), "Engaging Students in a Large Lecture: An Experiment using Sudoku Puzzles," *Journal of Statistics Education*, **22**(1), 1 - 20.

Cummiskey, K., Kuiper, S., and Sturdivant, R. (2012), "Using Classroom Data to Teach Students About Data Cleaning and Testing Assumptions," *Frontiers in Psychology*, **3**(354), 1 - 8.

Grantcharov, T. P., Bardram, L., Funch-Jensen, P., and Rosenberg, J. (2003), "Impact of Hand Dominance, Gender, and Experience with Computer Games on Performance in Virtual Reality Laparoscopy," *Surgical Endoscopy*, **17**(7), 1082 - 1085.

Jarrett, O. S., and Burnley, P. (2010), "Lessons on the Role of Fun/Playfulness from a Geology Undergraduate Summer Research Program," *Journal of Geoscience Education*, **58**(2), 110 - 120.

MacKay, R. J., and Oldford, W. (1994), "Stat 231 Course Notes Fall 1994," Waterloo: University of Waterloo.

Pfaff, T. J., and Weinberg, A. (2009), "Do Hands-On Activities Increase Student Understanding?: A Case Study," *Journal of Statistics Education*, **17**(3), 1 - 34.

Schwartz, T. A. (2013), "Teaching Principles of One-Way Analysis of Variance Using M&M's Candy," *Journal of Statistics Education*, **21**(1), 1 - 14.

Snee, R. D. (1993), "What's Missing in Statistical Education?," *The American Statistician*, **47**(2), 149 - 154.

Sullivan M. (2004), *Statistics: Informed Decisions Using Data*. Upper Saddle River, New Jersey: Prentice Hall, 2004 pg. 3 – 9.

Tchantchaleishvili, V., and Myers, O. (2010), "Left-Handedness - A Handicap for Training in Surgery?," *Journal of Surgical Education*, **67**(4), 233 - 236.

Wild, C. J., and Pfannkuch, M. (1999), "Statistical Thinking in Empirical Enquiry," *International Statistical Review*, **67**(3) 223 – 265.

Amy S. Nowacki
Department of Quantitative Health Sciences
Cleveland Clinic and Cleveland Clinic Lerner College of Medicine
9500 Euclid Ave. / JJN3 - 01
Cleveland, OH 44195
Email: nowacka@ccf.org
Phone: 216 – 444 – 3773