Doi: 10.14210/contrapontos.v16n3.p534-540

DEVELOPING AN INTUITIVE AND PROBABILISTIC UNDERSTANDING OF A=0.05 AND A=0.01 STATISTICAL SIGNIFICANCE LEVELS

DESENVOLVIMENTO DE UMA COMPREENSÃO INTUITIVA E PROBABILÍSTICA DE NÍVEIS DE SIGNIFICÂNCIA ESTATÍSTICA A = 0,05 E A = 0,01

DESARROLLO DE UNA COMPRENSIÓN INTUITIVA Y PROBABILÍSTICA DE NIVELES DE SIGNIFICANCIA ESTADÍSTICA A = 0,05 Y A = 0,01

Mehmet Türegün[1]

[1] *PhD in Mathematics Education from the University of Oklahoma. Lecturer at the Adrian Dominican School of Education - Barry University - Miami Shores – FL - USA.*

Abstract: The purpose of this article is to describe a modeling approach to help students develop a conceptual understanding of statistical significance levels, through discovery based on their own observations as they play the role of data detectives. I wish I could claim this clever idea as my own original thought, but in fact, Eckert (1994) outlined a similar activity. A few years later, Roxy Peck demonstrated a slightly different version of this activity in an Advanced Placement Statistics meeting. I have been using various versions of this modeling approach over the years, in my various statistics and quantitative inquiry courses at undergraduate and graduate levels.

Keywords: Teaching statistics; Modeling statistical significance; Data-based evidence.

Resumo: O objetivo deste artigo é descrever uma abordagem de modelagem para ajudar os alunos a desenvolver uma compreensão conceitual dos níveis de significância estatística por meio de descoberta com base em suas próprias observações quando eles desempenham o papel de detetives de dados. Gostaria de poder promover esta ideia inteligente como minha, mas, na verdade, Eckert (1994) delineou uma atividade similar. Alguns anos mais tarde Roxy Peck demonstrou uma versão ligeiramente diferente dessa atividade em um encontro sobre Métodos Estatísticos Avançados. Tenho utilizado várias versões desta abordagem de modelagem ao longo dos anos em meus vários cursos de estatísticas e pesquisa quantitativa nos níveis de graduação e pós-graduação.

Palavras-chave: Ensino de estatística; Modelagem de significância estatística; Evidência fornecida por dados.

Resumen: El objetivo de este artículo es describir un abordaje de modelado para ayudar a los alumnos a desarrollar una comprensión conceptual de los niveles de significancia estadística por medio del descubrimiento en base a sus propias observaciones cuando desempeñan el rol de detectives de datos. Me gustaría poder promover esta idea inteligente como mía, pero la verdad es que

Eckert (1994) delineó una actividad similar. Algunos años más tarde Roxy Peck demostró una versión ligeramente distinta de esa actividad en un encuentro sobre Métodos Estadísticos Avanzados. A lo largo de los años he utilizado diversas versiones de este abordaje de modelado en mis varios cursos de estadísticas e investigación cuantitativa en los niveles de grado y posgrado.

Palabras clave: Enseñanza de estadística; Modelado de significancia estadística; Evidencia proporcionada por datos.

INTRODUCTION

The increasing importance of collecting and organizing data, and exploring chance, are described repeatedly in the National Council of Teachers of Mathematics (NCTM) standards documents (NCTM, 2000; 1995; 1989) as major goals and trademarks of statistical activity across the school years. There is also a greater focus than ever before on using statistical methods to describe, analyze, evaluate and make decisions, while creating experimental and theoretical models of situations involving probabilities for grades 5-8 (Watson, 1998). Some of the most difficult inferential statistical concepts for teachers to teach, and for students to understand, are sampling distributions, standard error of the mean, statistical significance, and the logic behind hypothesis testing (Ryan, 2006). In the age of readily available computer simulations and applets, one can certainly use various applets available to demonstrate these ideas via simulations. However, starting with a hands-on simulation or model, and involving students in the process, has numerous intrinsic values, and is highly recommended by the statistics education research community (Chance & Rossman, 2006). Not only do hands-on models give students a sense of ownership, but they also improve conceptual understanding (Türegün, 2014). Hence, the use of a computer simulation has a less chance of becoming a meaningless activity using technology.

BACKGROUND

Typically, most classroom discussions on hypothesis testing involve statistical significance levels of $\alpha=0.05$ and/or $\alpha=0.01$, which are to be compared to the calculated *p*-values in order to decide whether or not to reject a null hypothesis. In their statistical diversions piece, Petocz & Sowey (2008) challenged teachers to come up with practical ideas for promoting statistical literacy. One of the questions they posed was as follows: Where do these "almost iconic" numerical values come from? When questioned as to where these significance levels of 0.05 and 0.01 come from, teachers generally refer to them as commonly used conventions. A few teachers might mention that they are commonly agreed upon definitions of chance occurrence. In other words, an event occurring at a rate of 1 out of twenty, or 1 out of a hundred, is likely to have happened by chance. Generally speaking, the discussion hinges on the use of significance levels $\alpha=0.05$ or $\alpha=0.01$ as conventional wisdom, and seeks to explain the reasons behind the use of these two significance levels. However, it falls short of developing a conceptualization of these significance levels.

What follows is a description of a modeling activity that can be used to help students develop a conceptual understanding of statistical significance. Two decks of playing cards with identical face designs are needed to carry out the activity. A twenty-dollar bill may also be used, at the end of the activity, to order a pizza for the class to enjoy while reflecting on the activity.

IMPLEMENTATION OF THE MODELING PROCESS

The modeling activity can be viewed as a three-stage process. The first stage is the prepping stage and takes place before entering the classroom. The teacher can start by taking the cards with identical design faces out of the boxes in which they were originally packaged. Since the jokers will not be used, take out the jokers from each deck, and set them aside. Arrange the decks so that each deck consists of only red or only black cards. Put the decks, now consisting of the same color, back into their original boxes. You now have a deck consisting of only black cards in one box, and a deck consisting of only red cards in the other box. Be sure to have an inconspicuous way of identifying which deck has all red cards, and which deck has all black cards.

The second stage is the romancing stage, which takes place in the classroom. You can start your class by announcing that you feel lucky and are in the mood for a little game of chance. Take out the $20 bill, show it to the students, and lay it out on the table. Take out the two decks of cards still in their original boxes, and ask the students to choose which deck they want to use. Make an offer of $20 to the first student who draws either a red or a black card from the deck, which depends on your inconspicuous mark to identify the deck color.

The final stage is the delivery stage. Take the cards out of the box, and ask a student to shuffle them. You may want to talk to the student while they are shuffling, to distract them from noticing that all the cards in the deck are the same color. Now, you have a nicely shuffled all black or all red deck of cards. Repeat the offer of $20 to the first student who draws a red card from the deck, if the chosen deck consists of all black cards, or vice versa. Let us assume that the deck with all black cards is being used. At this point, to add to the suspense and make the game more convincing, mention that to make it fair for all students, random assignment must be used to determine who is going to draw a card from the deck first. It is important to use random assignment if all the students are to have an equal chance of receiving the $20.

CONCEPTUALIZATION OF STATISTICAL SIGNIFICANCE LEVELS

Proceed with the first draw. To the disappointment of the first student, the card drawn is not a red card. You mention, at this point, that the chances of the first student drawing a red card was 50-50, or the probability was 26/52=1/2=0.5. Put the black card drawn back into the all black deck, and ask the next randomly assigned student to draw the next card. As you repeat this process, usually after the fourth or the fifth draw you start hearing students' murmurs. At this point, they begin to grow suspicious that something is not right, and something other than chance is at work. This is a good point to direct the discussion to the following probabilities of drawing a black card for several successive draws. The probabilities of drawing a black card successively are given as follows:

First draw: (1/2)

Second draw: (1/2)(1/2)=1/4=0.25

Third draw: (1/2)(1/2)(1/2)=1/8=0.125

Fourth draw: (1/2)(1/2)(1/2)(1/2)=1/16=0.0625 (*growing suspicious?*)

Fifth draw: (1/2)(1/2)(1/2)(1/2)(1/2)=1/32=0.0312 5

Sixth draw: (1/2)(1/2)(1/2)(1/2)(1/2)(1/2)=1/64=0.015625

Seventh draw: (1/2)(1/2)(1/2)(1/2)(1/2)(1/2)(1/2)=1/128=0.008

ISSN: 1984-7114

Based on these probabilities, you can form a group discussion on the following questions: Is what we have observed likely to occur by chance? In the absence of an underlying pattern in the population, what is the probability of obtaining a random sample like the one we have observed from the population? The subsequent discussions, sometimes with a little guidance, usually lead to the conclusion that if the probability is small, then we suspect that there may be some underlying pattern in the population. Even though what one may consider to be small is relative, in this activity, it almost invariably turns out to be between 1/16=0.0625 and 1/32=0.03125. Many scientific disciplines use 5 percent (.05) as the dividing line between small and not small when deciding whether or not an observed result is statistically significant.

EXTENSIONS AND IMPLICATIONS FOR TEACHING

In order to establish a tie to state a null hypothesis, a discussion can be started on the assumptions made by the students regarding the nature of the particular deck of cards used. It is important to point out the essence of the logic of significance testing. Start with a hypothesis (i.e., that the deck is fair), collect sample data (i.e., the successive draws), and ask whether the results observed would be surprising if the hypothesis were true. When the answers indicate that the results would indeed be surprising, reject the initial hypothesis (i.e., conclude that the deck is not fair). You may ask students to take a few minutes to write down this reasoning process in their own words, while remembering how the results were only unusual based on the underlying initial hypothesis of having a fair deck of cards. Technically speaking, even though this modeling activity may not be an accurate representation in the standard hypothesis testing sense, the activity is useful in demonstrating the conceptual understanding of statistical significance and significance levels, and motivates the students to think about probabilities and the reasoning behind hypothesis testing. In closing, best of luck to you all in using this modeling activity, and, most importantly, have fun.

REFERENCES

Chance B. L., & Rossman, A. J. (2006). **Using simulation to teach and learn statistics**. Paper presented at the 7th International Conference on Teaching Statistics (ICOTS-7), Salvador, Bahai, Brazil.

Eckert, S. (1994). Teaching hypothesis testing with playing cards: A demonstration. **Journal of Statistics Education**, *2*(1).

National Council of Teachers of Mathematics. (1989). **Curriculum and evaluation standards for school mathematics**. Reston, VA: Author.

National Council of Teachers of Mathematics. (1995). **Assessment standards for school mathematics**. Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). **Principles and standards for school mathematics**. Reston, VA: Author.

Petocz, P., & Sowey, E. (2008). Statistical diversions. **Teaching Statistics**, 30(1), 29-32.

Ryan, R. S. (2006). A hands-on exercise improves understanding of the standard error of the mean. **Teaching of Psychology**, 33(3), 180-183.

Türegün, M. (2014). A four-pillar design to improve the quality of statistical reasoning and thinking in higher education. **The Online Journal of Quality in Higher Education**, 1(1), 1- 8.

Watson, J. M. (1998). Professional development for teachers of probability and statistics: Into an era of technology. **International Statistical Review**, 66(3), 271-289.

Article received on: 07/07/2016

Approved on: 11/30/2016

Endereço para correspondência:

Mehmet Türegün. 5200 N. Ocean Blvd. #703, Lauderdale By The Sea, FL 33308.
E-mail: mturegun@barry.edu