THE "COMPLEAT" APPLIED STATISTICIAN

Donald Bentley
Pomona College, United States of America
dbentley@pomona.edu

In March of 2007, Discovery Channel released the controversial documentary, "The Lost Tomb of Jesus," which asserts that a first century tomb in Jerusalem contained the remains of Jesus of Nazareth and members of his family. Included in the documentary is an interview with a statistician who calculated the odds in support of this assertion at 600 to one. As skepticism intensified among biblical scholars, archaeologists, and the Christian community, support of the assertion relied more heavily upon the statistical results. Using the above example, this paper discusses the responsibilities of an applied statistician on a research team: becoming familiar with the subject of the research, scrutinizing every assumption used in the analyses, validating all data, and being sensitive to consequences the statistical results might have on society.

INTRODUCTION

Over the past half century, major changes have taken place in statistics education and in particular in introductory courses. During the early years, courses focused on the analysis of data where the "data" consisted of sets of numbers without context, but which were computationally manageable with crude mechanical calculators. The next stage introduced a context to fabricated data, making it appear "realistic" but still allowing for ease of computation. Thanks to more powerful calculators and the introduction of computers, instructors became able to provide real data sets in a real context for the students to analyze (for examples see Witmer, 1992; Rossman & Chance, 2001).

Currently, we are finding success in motivating students with the inclusion of student projects in which they participate in the collection of the real data (for examples see Schaeffer et al., 2004). The majority of these projects are activities designed to expose the students to a specific statistical methodology or concept. In these, all students work on the same project, having ownership of their data but not of the experiment or the context.

Some introductory courses now require a term project in which students design their own individual studies, gather their own data, and present their own analyses. This total process not only provides each student with a better understanding of the context of the issues the study is addressing, but also provides a sense of ownership in the project which motivates enthusiasm, and removes the drudgery of sterile "number crunching" which was so tedious for students not so many decades ago.

While the changes in statistics education mentioned above have had a positive effect on students' attitudes about statistics as well as their proficiency in statistical reasoning and understanding of statistical concepts, this has only affected students in those courses where these changes have been implemented. However, there remains a perpetual debate in many academic institutions which precludes a large number of students from the opportunity to be exposed to these innovative teaching methods. This debate is between *mathematics/statistics departments* and *service departments* which are those of academic disciplines in which statistics is considered to be a standard tool in the same way mathematics is a tool in physics. These include departments of such disciplines as economics, sociology, psychology, and biology which often offer their own introductory statistics courses. Their justification for teaching their own courses is that statisticians, and consequently the service courses statisticians teach, don't teach "the standard material" (i.e. the statistical methodology) their students need. And further, they don't provide a context for the material with which their students are familiar and in which they are comfortable.

Statisticians, on the other hand, argue that the faculties of these service departments are not properly trained in statistics and the quality of the statistics presented in the service department courses is questionable at best. This impression is frequently confirmed by student performances in advanced courses for which these service courses serve as a prerequisite. But there is another side to this coin. Perhaps we statisticians can learn from the complaint of the service departments which I believe suggests additional concepts which should be included in the introductory statistics

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

courses, both those serving as service courses for other disciplines and those first courses for students intending to pursue statistics as a career. And in creatively teaching these concepts we might make our courses more appropriate and acceptable to service departments.

The purpose of this paper is to tell a story. It is a true story which involves a statistician and the analysis of data which suggests this additional material in order to expose students to the complete role that an applied statistician is obligated to perform as both a member of a study's research team and also as a member of society. The story is of a statistician and his analysis of a data set in which the data and the results of his analysis became extremely controversial after the results were released to the general public. It is a story which I believe points out several important responsibilities of the applied statistician which we, as teachers of statistics, seldom address in our teaching of the introductory courses. I now realize I have subconsciously assumed these responsibilities during my career as an applied statistician while consulting for biomedical companies throughout my teaching career, and more recently in applying statistical methodology to biblical archaeology on excavations in Israel. But these responsibilities were never more vividly impressed upon me than through this story which is of my involvement in a project Allan Rossman convinced me to accept in the spring of 2007, and which subsequently consumed my life for the major portion of a year.

ROLE OF THE "COMPLEAT" APPLIED STATISTICIAN

Perhaps the foundation for this story was laid in 1996 when I was invited to write a chapter in a book titled *Education in a Research University* (Bentley, 1996) honoring Jerry Lieberman on his retirement as Provost of Stanford University, and shortly before his death from ALS (Lou Gehrig disease). The topic of my chapter was a lecture I had developed for the first day of class in all my introductory statistics courses. The purpose of the lecture was to attempt to introduce the students to the role of an applied statistician as being something more than just applying a collection of memorized formulae to data. Instead, the role is that of a member of a research team to be involved throughout the total study. The paper identifies five areas of a research project in which the statistician, to be effective, must be intimately involved. These five areas are: a) forming the questions, b) designing the experiment, c) gathering the data, d) analyzing the data, and e) communicating the results. The importance of the role the statistician plays at each of these points was illustrated by example.

Several years later, Lincoln Moses (my mentor whom I consider to have been one of the top three applied statisticians of the 20th century) happened to be visiting. Somehow we wandered into a conversation as to the role of statistics and the statistician in scientific research. Lincoln summarized his opinion as follows: "Mathematics is known as the Queen of the Sciences. *Statistics is the Umpire of the Sciences.*" (Personal Communication). In other words, Lincoln believed it to be the role of applied statisticians in every research project in which they are involved, to make sure the research team abides by the rules appropriate for that discipline. And while I feel my list of the five areas where an applied statistician must be involved provides some guidance for the statistician to perform as an umpire, I have only recently realized that the list needs augmenting with at least two more responsibilities which I describe below. Their importance will then be illustrated with the story that follows.

A major aspect of applied statistics which is seldom if ever emphasized in statistics courses is that research projects typically have two sets of players, often with conflicting interests. One set is what might be called the statisticians' clients, those conducting the research and generating the data that are to be analyzed. This is usually the group that has initiated the services of the statistician. The second set would be the consumers, those who will be affected by the results of the analysis. For example, in a clinical study of an investigational drug the pharmaceutical firm becomes the statistician's client and all potential victims of the disease for which the drug might be an effective treatment become the consumers. Students need to be made aware that it is a responsibility of the statistician, as umpire, to enforce the rules to make sure that not only the statistical analysis of the data but also the total investigation is conducted according to the rules so as to protect the interests of both the client and the consumer. This is not only a statistician's scientific responsibility, but it is also an ethical responsibility. And often it is difficult

for a statistician to remain a neutral umpire as enthusiasm generated within the research team of which the statistician is a member has a tendency to become contagious.

The second added responsibility follows from the need of statistical reasoning to be applied at each phase of a project, not just during the analysis of the data. The applied statistician needs to be involved from the very beginning, from the forming of the questions the research team will attempt to answer. It is the responsibility of the statistician to make certain everyone understands the questions and is in agreement that they are the appropriate questions for the project. Moreover, equally as important is that the statistician must make sure everyone is in agreement with the assumptions behind the questions. This includes making sure these assumptions form a consistent set. There should be ample evidence that they can be reasonably accepted not only by the complete research team, but by the community of scholars in the field of application. The statistician, though not an expert in the substantive field, should still feel comfortable with the arguments being used by the research team members to justify the assumptions. And the assumptions should not preclude being able to answer the agreed-upon questions; they must be consistent with the questions. Keep in mind that a good umpire is not required to be highly skilled in playing the game, but must have a good knowledge of the rules by which the game is to be played and have had enough exposure to the game to be able to detect when the rules are being violated (Bentley, 2008). Therefore applied statisticians, in each research project in which they are involved, must be familiar with all the rules the researchers are to follow. They must have knowledge in the discipline of the study. A statistician who does not feel comfortable with his or her knowledge of the specific area of a research project must either devote the necessary time to become proficient in that area, or else decline the invitation to become involved in the project. Most applied statisticians find this excuse to extend their knowledge a benefit of the discipline.

THE JESUS FAMILY TOMB STORY

This particular story begins the morning of Monday, February 26, 2007 at a press conference called by The Discovery Channel, and held in the New York City Public Library. (To view the press conference, link to http://dsc.discovery.com\converge\tomb\tomb.html and choose "Watch Press Conference.") It was attended by reporters from the major news agencies from around the world. The news conference began with the following greeting:

> Good morning. Thank you for joining us for this very exciting announcement. I'm Jane Ruth, president and general manager of the Discovery Channel, and you are joining us here for what might be one of the most important archaeological finds in human history. In the hills of Jerusalem archaeologists have discovered a tomb, a 2000 year old tomb which contains significant forensic evidence, and some potentially historic consequences. [. . .]
>
> Today we are bringing you new scientific analysis of a tomb that was first discovered in 1980. [. . .] In this particular tomb, there were 10 limestone boxes called ossuaries, a common form of burial in the first century. Six of these boxes were inscribed with names, very important names. Given this, and other forensic details that we are about to reveal to you today, we believe that there is compelling data that these tombs may have contained the remains of Jesus of Nazareth and several members of his family. [. . .]
>
> Our documentary, "The Lost Tomb of Jesus," is produced by Academy Award winning filmmaker James Cameron [producer of the movies *Titanic* and *Avatar*], who barely needs any introduction, and is directed by Emmy Award winning documentarian Simcha Jacobovici, both of whom you will hear from in a minute.

Ms. Ruth continued by introducing the panel of experts who had worked on the project, and who were present on the dais to answer questions of the press. This group included two biblical scholars, an archaeologist, and a statistician. She then turned the podium over to James Cameron. But before considering his introductory remarks, I think it appropriate to review those made by Ms. Ruth from the perspective of an applied statistician in the role of the umpire for the research.

From the introduction of the panel, we know the statistician is a member of the research team. As such, the Discovery Channel becomes his client and the documentary, "The Lost Tomb of Jesus." is the product being developed. The consumers will be those people who will be affected

by distribution of the product. For "The Lost Tomb of Jesus" documentary, this group consists not only of the potential audience that will view the documentary, but also all of those who might ultimately be affected by the distribution of the results presented in the documentary. These results are the claim of the discovery of a tomb which contained ossuaries which possibly contained the bones of Jesus of Nazareth and Mary Magdalene. The immediate consequences of the existence of such ossuaries would be the disproving of the physical resurrection of Jesus of Nazareth. Such a result would not only have great historical consequences, as suggested by Ms. Ruth, but also tremendous theological consequences because it would contradict a tenet basic in the faith of billions of Christians around the world.

A possible framework for the statistical decision process would be that of a hypothesis test, in which the alternative hypothesis becomes that this particular tomb (Talpiyot) is that of the Jesus family, and the specific ossuary that of Jesus of Nazareth. One can then discuss the consequences (losses) associated with both type one and type two errors. For the client (Discovery Channel), a type two error of not concluding the tomb to be that of the Jesus family would have the financial consequences of loss of investment in the project. However, for the consumer which includes not only biblical scholars, but those Christians for whom the physical resurrection is central to their faith, the consequences of a type one error would be immense.

Following Ms. Ruth's comments, James Cameron took the podium and began by stating:

> Now I've never doubted that there was a historical Jesus, that he walked the earth 2000 years ago. But the simple fact is that there has never been a shred of physical archaeological evidence to support that fact until right now. What this film and the investigation that the film shows is able to bring to light is, for the first time, tangible physical archaeological and in some cases forensic evidence; forensic evidence that can be analyzed scientifically that found it to be, to a layman's eye because I quickly profess that I'm not an archaeologist, I'm not a biblical scholar. But to a layman's eye it seemed pretty darn compelling. And as a documentary film maker, I was very, very attracted to the story. I said, "I think that literally this is the biggest archaeological story of the century."

Cameron concluded his remarks by turning the podium over to Jacobovici who began, "It is somewhat surreal to be in the New York Public Library, the lights on, the media here, and to know that underneath that felt are possibly the coffins, the bone boxes of Jesus of Nazareth and Mary Magdalene."

After further discussion, he described his qualifications. "My expertise is in investigative journalism, and this is what I do for a living; and it is a skill set that many people in this room share. I'm not an archaeologist, I'm not a DNA expert, I'm not a statistician, I am a filmmaker and a journalist." Jacobovichi then described the data that was used in the documentary to support his hypothesis that this tomb, Talpyiot, was in fact that of the Jesus family. Of primary importance was the interpretation of the inscription Mariamneou [eta] Mara which was found on the ossuary Jacobovici identified as that belonging to Mary Magdalene. After describing the collection of ossuaries and the inscribed names, and their relation to those names known of members of the Jesus family, he proceeded with,

> We asked the archaeologists, when you say these are common names and statistically insignificant, have you ever spoken to a statistician. And the answer universally was "No". So we did what good journalists do. We went to statisticians. And the range was two million to one in favor of this tomb, to the low end of 600 to one in favor of the tomb. We went with Professor Feueverger's numbers which were the most conservative of the several studies that we initiated.

Following Jacobavici's remarks the panel accepted questions from the news media. One in particular was from a reporter from Time Magazine who questioned the influence of the interpretation of the Mariamneou inscription as "one that really rigs the odds." The statistician responded to this question with the following:

> The obvious needs to be stated, that I'm not a biblical scholar, I'm not a historical scholar, I'm just a numbers guy, a numbers guy for the project. [. . .] As a statistician, I did the calculations based on assumptions given to me by the subject matter experts, in this case historical biblical scholars. One of the assumptions that really helped to drive the calculations is that in fact Mariamneou [eta] Mara is a very highly appropriate appellation

> for a particular individual that we are talking about here, and Jose the extraordinary unusualness of this. When you factor in the unusualness of the names and the particular configuration that you get, and then you try and ask yourself, go back in time, and ask yourself what is known about the demographics of the era, you can try to work out the odds of, what are the odds that there might have been another family whose tomb this might have been
>
> The numbers you get depend on the kind of assumptions you put into the analysis. As a statistician it's not my job to take responsibility for the assumptions, but it's my job to do the calculations as carefully as I can. [. . .] And based on the assumptions that I had, and I've worked this through as carefully as I could, we're seeing numbers that are actually such that they would make you think, you should stop and pause, and this might be it. The numbers could range, depends on which assumptions you're willing to play with, they could range from one in a hundred to one in a thousand against there being some other family having this particular configuration of names. But that's roughly where I'm coming from.

THE DATA AND ANALYSIS

On Sunday evening of March 4th, 2007, the documentary *The Lost Tomb of Jesus* was released for general viewing on the Discovery Channel. Immediately it was hotly criticized by many members of the biblical scholarship and archaeology communities. One of the major complaints leveled at the documentary was its lack of a comprehensive peer review. As time passed, the Discovery team attempted to deflect the criticism from these communities by placing more emphasis on the results of the statistical analysis; that the odds were at least 600 to one in favor of the tomb being that of the Jesus family.

At the same time that this debate was heating up, Allan Rossman, serving as program chair of the 2007 Joint Statistics Meetings (JSM), was arranging two special sessions reserved for what were designated as "Late-Breaking Sessions," dedicated to cover important topics emerging close in time to the meetings. For one he invited the statistician for the documentary to present his analysis of the Jesus tomb data. He then contacted me and asked if I would be willing to be one of the three discussants.

Immediately I knew I should decline the request, I should avoid getting involved with this controversy. But after the initial shock wore off, I questioned who he might have for discussants and I realized they would likely be statisticians (tending more toward mathematical than applied) with no experience in biblical scholarship or archaeology. I also recognized my background of a seminary education and my experience in applying statistical methodology to archaeological excavations in Israel would give me an advantage in serving as an umpire when evaluating the research. Following the presentation at the JSM it was announced that the statistician's paper was to be published in the Annals of Applied Statistics, along with discussion.

My discussion did not address the methodology of the statistical analysis, as I left that for other discussants. Instead I focused on points raised in the original news conference which were reinforced in the paper. In particular, in the news conference the statistician claimed he "did the calculations based on assumptions given to me by the subject matter experts, in this case historical biblical scholars." In his paper in AOAS (see footnote 33) Feuerverger (2008) lists Jacobovici as the source for seven of the nine majors assumptions upon which he based his analysis. In particular is assumption A.7 which is that, "The inscription Mariamenou [eta] Mara . . . represents the most appropriate specific appellation for Mary Magdalene among those known." Yet in the news conference, Jacobovici specifically stated, ". . . I'm not an archaeologist, I'm not a DNA expert, I'm not a statistician, I am a filmmaker and a journalist." Jacobovici specifically excludes himself from being a qualified expert, yet the statistician bases the analysis provided by this non-expert. And one of these assumptions was A.7 which the statistician, in the news conference, stated, "One of the assumptions that really helped to drive the calculations is that in fact Mariamneou [eta] Mara is a very highly appropriate appellation . . ." for Mary Magdalene.

In the documentary, (and also in the book *The Jesus Family Tomb*, Jacobovici & Pellegrino, 2007), Jacobovici attributes this assumption to a discussion with François Bovon, Professor of History of Religion at Harvard Divinity School whose specialty is New Testament

and early Christian literature. However, shortly after release of the documentary Bovon made a statement through the Society of Biblical Literature (Bovon, 2007) stating he does "not believe Mariamne is the real name of Mary of Magdalene." He points out that when he was questioned by Jacobovici and his team, "the questions were directed toward the *Acts of Philip*, a non-canonized fourth century text, and the role of Mariamne in this text. I was not informed of the whole program and the orientation of the script." In other words, Jacobovici's assumption is a misinterpretation of Bovon's remarks. The statistician, as an umpire of the science, should have confirmed this assumption with Bovon, the source, and not accepted it from one of the research team who had a vested interest in the script and was by his own admission, not an expert in the disciplines involved.

THE REST OF THE STORY

In the particular issue of the Annals of Applied Statistic which included the statistician's paper and the subsequent discussions, the author presented a rejoinder, in which he laudably acknowledged that Bovon's comments "had inadvertently been misinterpreted by Jacobovici – were conveyed to me out of context" (Feuerverger, 2008, pp. 99-112). He therefore no longer considered assumption A.7 adequately justified, and consequently continued, "In particular, this means that we cannot . . . say that the Talpiyot find is statistically significant in a meaningful way." At this point, he had assumed the proper role of a "compleat" applied statistician But this occurred only after he went beyond just performing statistical analyses of the data, only when he questioned the validity of the assumptions that were basic to the formation of the hypotheses being considered, only when he assumed the role of umpire that his performance led to the scientifically and ethically correct conclusion.

REFERENCES

Bentley, D. (1996). My First Days' Lectures: Past and Present. In K. Arrow, R. Cottle, B. Eaves, & I. Olkin, (Eds.), *Education in a Research University.* Stanford, California: Stanford University Press.

Bentley, D. (2008). Discusssion. *Annals of Applied Statistics, 2*(1), 66-73.

Bovon, F. (2007). The tomb of Jesus. *SBL Forum*. http://www.sbl-site.org/publications/article.aspx?articleId=656.

Feuerverger, A. (2008). Statistical Analysis of an Archeological Find. *Annals of Applied Statistic, 2*(1), 3-54.

Feuerverger, A. (2008). Rejoinder. *Annals of Applied Statistics, 2*(1), 99-112.

Jacobovicci, S., & Pellegrino, C (2007). *The Jesus Family Tomb: The Discovery, the Investigation, and the Evidence That Could Change History.* San Francisco: HarperCollins Publishers.

Rossman, A., & Chance, B. (2001). *Workshop Statistics: Discovery with Data* (2nd edition). Emeryville, California: Key College Publishing.

Schaeffer, R., Gnanadesikan, M., Watkins, A., & Witmer, J. (2004). *Activity Based Statistics*. New York: Springer-Verlag.

Witmer, J. (1992). *Data Analysis: An Introduction.* Englewood Cliffs, New Jersey: Prentice-Hall.