Cecil Thompson

Omega Graduate School

COM 968-52

Prof. Sean Taleday, Ed.D

Prof. Richard Nti, M.S.

March 19, 2024

Assignment #2 – Developmental Readings:

1. Create Developmental Readings from seminal sources and scholarly peer - reviewed

journal articles. Review instructions for Assignment #3, the course essential elements,

and course readings to identify selections of books and journals to create entries.

1. Refer to the " Student Guide to Developmental Readings " in the General Helps

folder for updated information on sample comments , the grading rubric , and key

definitions related to developmental readings.

**Source 1**:

Hardwickes, Tom E; Goodman, Steven N. (2020, Oct). PloS One; *San Francisco*. Vol. 15(10):

e239598. Doi: 10: 1371/journal.pone. 0239598. [Method]

**Comment 1**:

**Quote/Paraphrase**: Biomedical literature is usually laden with statistical analyses that give scientists in that field the means by which to make scientific claims. But due to frequent misunderstanding and misuse of scientific data, such scientific research, at times, fails to achieve the researcher’s intended goal. The reason why biomedical scientists employ statistics into their research is not only to quantify data in order to validate their opinions, but to present the opinions in a manner that will be acceptable to others in the biomedical science community, as well. One of the ways in which such research can be made available to them is when it is published by well-respected media in the field. Many researchers in this and other fields are not trained to conduct their own statistical work, and therefore, some may need to rely on experts who are specialists in the field of statistics (npn).

**Essential Element**: This has to do with the general nature of statistics and the family of discipline to which it belongs.

**Additive/Variant**: Statistics is a subordinate discipline within the much broader discipline of mathematics. Since mathematics is such a large family of disciplines dealing with its practical and theoretical applications to known and unknown quantities, together with activities involving mensuration, it seems grammatically apt that in English, as in a number of other modern languages, it would take the form of what is functionally a collective noun with an implied plurality in its “s”-ending. Therefore, although the syntactical agreement between number and case remains singular, many specialized sub-disciplines, such as statistics, would still huddle together underneath the huge umbrella of mathematics. Definitively, statistics may be said to be that branch of mathematics that is used to convert data into quantifiable form for the purpose of making predictions about the future state or behavior of phenomena. In this article, the concern is about the extent of the usefulness of statistics in the field of biomedical science.

**Contextualization**: As a teacher of mathematics, I have taught statistics from the very elementary level up. For example, I have taught central tendency, percentages, percentile, “descriptive statistics,” distribution (normal and skewed), combination, permutation (replacement set, non-replacement set), binomial expansion, Bernoulli’s experiment, correlation, exponential growth, regression, etc. Initially, I taught standard deviation through the tedium of the conventional timeworn tabular algorithmic system until NY math teachers were forced to give it up for the graphing calculator. But my resistance to the way technology had displacing “good” old-fashion mathematical algorithms, and my failure to see it as part of the body of applied mathematics left me well behind the eight ball when it comes to computer application in statistics. Rather than tapping data into a trusty computer, I still had the tendency to subject myself to the error-prone tedium of the Pearson r formula and algorithm when trying to figure out the relationship between two variables.

**Comment 2:**

Q**uote/Paraphrase**: A ***Sample*** of journals is collectedfrom a list of 228 “Web of science subject categories”. Sixty-eight subject-areas of these belonged to the area of biomedical research. Some were ranked at the top because of their impact in each of the various sciences being studied. Five journals were selected from the top of each of the subject areas. Prior to this selection, Steven Goodman (2020). one of the authors of this article, and his colleagues had conducted a survey of journals with biomedical subject areas. Sixty-eight of them were added to their relevant subject areas on the list. In the end, journals that carried articles in multiple subject areas were excised, so that, the total selection for the sample was 364 journals (npn).

**Essential Element**: In order to mitigate bias and increase the probability that the sample will be reflective of the population, special attention has been paid to the manner in which the sample is collected as well as the sample-size.

**Additive/Variant**: One has to pay close attention to how the sample is collected and whether the sample-size is adequate to give the best possible representation of the population in the type of research that is being conducted. The smaller the sample, the less certain that it will be a true representation of the population. In some scholarly research, a sample size of thirty is considered to be sufficient. My thinking is that this kind of research does need a more sizeable sample than 30, and so, the sample of 364 journals appears to be a good choice for this research.

**Contextualization**: As a teacher of mathematics, most of the statistical skills I have taught have been taught from textbook-type of fictional data through rigid algorithmic manipulations rather than by means of data entry into a graphing calculator or some other kind of software driven device. Since the emphasis, in the high school mathematics lesson, then, is placed on the imparting of concepts and skills, rather than on research, I was obliged to use sample sizes that had mostly about 8 or 10 participants. So, clearly the objective in the teaching of statistics under a high school mathematics curriculum places more emphasis on the imparting of skills and concepts than it does about research.

**Comment 3**:

**Quote/Paraphrase**: ***Method***. The digital survey instrument used (check: <https://osf.iodg9ws/>) was an adaptation of that used by Steven Goodman. The first (Q-1) of the 15 questions asked was: “Of original research articles with a quantitative component published in your journal, approximately what percentage has been statistically reviewed?” For those whose answer was less than 10%, the question was skipped. If the answer from the person representing a journal was greater than 10%, he would have to answer questions related to statistical review policies. A question concerning statistical competence was asked of all except those representing journals not requiring statistical reviews, or those stating that statistical aspects of articles are dealt with in peer-reviews or by editors. After the three questions about the characteristic of the journals were asked, each representative was asked to make a free-text comment. The complete survey instrument is available (check: <https://osf.io/dg9ws/>) together with précis versions of the questions and responses, as approved by Stanford IRB#42023 (npn).

**Essential Element**: This has to do with the screening and elimination of samples that do not fit the population identified for the research.

**Additive/Variant**: Clearly, the method is focused on eliminating the journals of which there were very little or no statistical elements in the articles they published. And so, the journals that were weeded out, through this method of research, were those that added no real weight to the body of the research of those journals involved in the quantitative research.

**Contextualization**: In my experience as a teacher of mathematics, viz., statistics, our mode of instruction has had precious little to do with sifting out irrelevant portion of the samples we use because I did not teach statistic as a form of applied mathematics. But now, in my current study of statistics, this approach will obviously be invaluable in how I conduct sampling during my scholarly research, going forward.

**Source 2:**

Bala, Jyoti. (2016 Nov). Contribution of SPSS in social sciences research. 7(6) (Special Issue)

*International Journal of Advance Research in Computer Science*.

[www.ijjarcs.inform](http://www.ijjarcs.inform).

**Comment 4**:

**Quote/Paraphrase**: In recent years, researchers have been accumulating a tremendous lot of knowledge from surveys and empirical processes. And now, in the wake of the burst of new advancements in information technology, computers have become invaluable to researchers in all types of research. Because of the rapid rate at which computer applications are being created, in the area of social research, scientists have been able to tap into what would ordinarily have been among the hidden, if not inaccessible, complexities of man’s nature and his social organizations. Consequently, social scientists have come to rely upon the computer for just about every aspect of research involving human society. Just like the scores of general-purpose codes and the compatible hardware that are being churned out of the heads of Silicone Valley coders, and elsewhere in the geek world, billions of lines of special purpose codes are also being etched into storage devices to advance the research objectives of scientists in their various specialized fields. The *Statistical Package for the Social Sciences (SPSS)*, although widely used by researchers in diverse fields, had rightly become the all-purpose specialized software for social scientists. The need for statisticians to demonstrate the facility to respond to difficult research questions about society by way of ponderous mathematical calculations, is fast becoming a thing of the past. The SPSS is now supplying answers to the most enigmatic questions about society with a couple of clicks on a keyboard (p. 250).

**Essential Element**: This is about the building of a statistical bridge across the chasm between the old approach of using mathematical formulas and algorithms to solve statistical problems by a few and the newer highspeed use of SPSS software, accessible to all.

**Additive/Variant**: *The Statistical Package for the Social Sciences* (SPSS) had been, for some time, perhaps a bigger deal than people such as I, living outside of the statistical loop, am able to fathom. But today, Microsoft Excel has developed to a point where it has tremendous statistical capabilities. Meantime, SPSS, having pushed the door of statistical software wide open, has led the way for the development of programs like PSPP. Now, it appears certain that newer and more advance statistical software are being written to help statisticians solve research problems without the need for highly specialized labor-intensive mathematical skills than those needed in decades past.

**Contextualization**: Over at least the last twenty-five years, school districts have been requiring that high school mathematics teachers use the TI-83 and TI-84 Calculators to teach logarithms, statistics, and other sub-sets of mathematical disciplines and functions that are programmed in it. School examinations prepared by the various Boards of Regents and Advance Placement boards have long required the use of the TI-83 for their examinations.

Today, teaching the use of the TI-83 or TI-84 have now become imperative for the teacher of mathematics, viz., statistics. Although the teaching of SPSS and PSPP have not yet become common instructional software tools in the statistical part of the high school mathematic curriculum, I believe it is just a matter of time.

**Comment 5:**

**Quote/Paraphrase**: The *Statistical Package for the Social Sciences* (SPSS) was first written in 1968, by Norman H. Nie, Dale H. Bent, and C. Hadlai Hull as a statistical research tool specifically for analytic work in the social sciences. SPSS has been upgraded three times in 47 years. It is a software program with numerous subroutines, capable of reading data from questionnaire. It always begins with definitions of variables, graduating to bivariate tests, correlations, nonparametric tests, and so on. Next, upon the heels of the identification of the variables, comes the quantification, or measurement of the variables. Thus, numbers are assigned to nominal variables, and the likes. Although the researcher need not be a mathematician to be able to use the SPSS effectively, he or she must be sufficiently facile with most of the statistical terminologies around which the software program revolves (p. 251).

**Essential Elements**: When did the *Statistical Package for the Social Sciences* (SPSS) first appear and what is its capability?

**Additive/Variant**: I still believe that, while researchers may be passing their statistical problems on to specialists who are armed with dependable software, with scores of statistical functions, the statistician, himself, should be armed with the pertinent back-up mathematical skills despite the availability of software such as SPSS, just in case. At any rate, SPSS turned out to have been something of a pioneer for other statistical software such as PSPP.

**Contextualization**: Many high school mathematics teachers have become so taken by the speed of the TI-83 and TI-84 technology, that they no longer teach the corresponding formulas and algorithm. There is good reason why I have been careful to teach both the manual mathematics, including derivation of formulas and attendant algorithms, along with the calculator skills. I suspect that as with the TI-83 and TI-84, there are also special skills one should learn in order to be able to use the *Statistical Package for the Social Sciences* at the highest level of proficiency. Also, I think that since a part of the job of the teacher of mathematics is to build a foundation that undergird any of the many the mathematics-dependent careers that student might take. For instance, if a high school graduate becomes a computer software developer for programs such as SPSS, he or she has to be able to do the math.

**Comment 6**:

**Quotes/Paraphrase**: How is SPSS serviceable to the statistician whose job it is to solve statistical problems for researcher? To use this software, the variables must, first, be **quantified cation**, then, the **scales of measurement**, ranging from “nominal”, “ordinal,” “interval,” to the “ratio scale” are itemized. Following this, the **Study Design** has to be determined. The first is called, “between-participants (unrelated) design,” and the second is, “within-the-participant (related) design.” The first has to do with a comparison between two or more groups. The second design has to do with using the same sample repeatedly to measure different things. In addition, the same thing could be measured repeatedly. Of course, if there are two or more independent variables, the study could be subject to a “mixed-design”—the two different variables may be used both as a *related* design as well as an *unrelated* design (p. 242).

**Essential Elements**: This article has to do with how SPSS is used in meeting the various needs of the statistical work in research as to whether the study-design it relates to the study-design.

**Additive/Variant**: In using SPSS, it is necessary for all types of variables to be converted into quantifiable form, even though those such as categorical, nominal, or ordinal variables do not naturally exist as “quantitative data” when they are collected. This article, in its service as a medium by which to describe the way this software technology perform statistical tasks, has unveiled for us a palpable grasp of what is meant by “study design”.

**Contextualization**: To illustrate the concept of what may be called an “unrelated study design”, 60 adults between the ages of twenty-five and thirty could constitute a sample selected at random. In the beginning, they would all be tested for COVID-19 and come out negative. Once done, the sample would be divided into two halves. The first (experimental half) group of thirty will be given the J & J vaccine. Meanwhile, the second (control) group will be given a placebo. This is an example of the “unrelated” or, the “between-participant” study-design. After two weeks, and an exit test result of 100% negative, the participants will be released.

To illustrate the second type of study-design known as the “related” or, “within-participant” study-design, a sample of thirty will be selected from a similar population. These would be domiciled in a semi-quarantined common hotel-living facility together. They would be expected to prepare their own meals so that they would not come in contact with any outsider once indoors. However, they would be allowed to continue all their usual outdoor activities while wearing a mask. After two weeks, they would again be tested. Once the test result is negative for the pandemic, the participants are release. The same group will be asked to remain in the living facility for two more weeks. This time they were asked to not wear masks for outdoor activities. In the end, they would again be tested. This time, 30% of them would test positive and 20% of those who are testing positive would wind up being hospitalized.

**Comment 7:**

**Quote/Paraphrase**: Here, a case is being made about how to set off the research as soon as the values are assigned to the variables. The data is loaded into the SPSS. This will be the most time-consuming part of the process. The cursor from the mouse is put on the “Menu”; it clicks on the variable of choice, followed by a click on “Run!” SPSS then reads the cases, does the analysis, and spits out the results. All the results are then sent to a dialogue box named, “SPSS Viewer.” “The SPSS is a desktop and larger computer-based quantitative analysis package . . . [it] can perform many data management and statistical [analytical]. . . tasks with base SPSS, a researcher can manage data with case-selection file reshaping and creating derived data. A metadata dictionary is stored with the data statistical analysis tasks that can be performed with the base package include the generation of descriptive statistics, bivariant statistics, prediction of numerical outcomes, and prediction of identifying groups. Although SPSS is very popular with social science researchers, its ease of unsend-add-on modules allows it to operate as a cross disciplinary software package” (p. 252).

**Essential Element**: How is the SPSS package be used along with the “SPSS Viewer”.

**Additive/Variant**: In this segment, instruction for the use of SPSS actually takes us through the software methodology for the standard research process. This segment begins with the “Menu” and continues to where there is a command to “Run”. I would think that since the results of the analyses that follow the command to run are sent to the dialogue box, SPSS Viewer, the program does probably have some way to provide the researcher or the statistician with a way of viewing the results so that he or she will be able to make some corrections. The fact that it does the two important functions of “data management” and “statistically analysis”, is just about everything that is needed.

**Contextualization**: I am not conversant with the SPSS, but I think that it ought to be a software that provides plenty of solutions through “descriptive statistics” along with all the alpha-numeric outputs. From my familiarity with the TI-83, I must applaud it for being a complete statistical tool that produces results in graphs, histograms, etc., as well as numeric results.

**Source 3**:

Andrews, Douglas M. (1999 Aug). The American statistician. *Journal of Statistics*

*Education,* Alexandria, 53(3), 291.

[Statistical Reasoning & Method for pedagogical purposes].

**Comment 8**:

**Quote/Paraphrase**: This article is a Douglas Andrews review of the book, *Statistical Reasoning and Methods* by Richard A. Johnson and Cam Wah Tsui. Andrews points out that the book is actually an attempt to tie together a statistical bundle of loose ends from an earlier book entitled, *Statistics: Principles and Methods*. The original book was written by Johnson and Bhattachary in 1987. Andrews et al, have, in this new book, written a follow-up text to illuminate the chapters on variance analysis, multiple regression, and non-parametrics. In it, they have compressed categorical inferences and inferences for one and two proportions into a single chapter dealing with analysis of “count-data”. He says that in this work, they are responding to a public outcry that statistics education should place emphasis on statistical reasoning over calculations, which can now be left to computers. The idea is that this version should include actual examples of learning data, student projects and exercises. But, since most of this material is lifted straight out of the previous text, this book has neither changed focus nor orientation (p. 290).

**Essential Element:** This segment questions whether the statistic text for statistic education should emphasize statistical reasoning over how to solve statistical problems.

**Additive/Variant**: The criticism, here, is being leveled mainly against what the reviewer sees as an unnecessary over-use of mathematical formulas and algorithms to solve statistics problems related to variance, multiple regression, and non-parametrics. The position is based on the fact that most solutions to statistical problems only require a few data entries. The problem is that when the computer is doubtful, how do we check or fix it without resorting to statistical formulae and algorithm? For that matter, both are necessary.

**Contextualization**: Some years ago, my PSA (prostate specific antigen) began to edge up. I looked at my *multiple myeloma* m-protein and PSA and thought that both of them seemed to have been moving up and down in synchronicity. I decided to treat the m-protein as an independent variable (x) and the PSA as the dependent variable (y). I did the calculations manually using the Pearson r—a massive formula. The coefficient turned out to be .86 (86%). I could not believe it. I felt I may have made a mistake, so I tried it on the computer for verification. The computer result was the same. I then used the Kapa/Lambda ratio from the light chain proteins on the myeloma cells as the x and the PSA as the y. The correlation coefficient was .85 (85%). And, once again, I tried the computer for verity. The result was identical. These were high correlations, and because of that, my doctors were able to make better decisions on the treatment choices.

**Comment 9:**

**Quote/Paraphrase**: Although the authors have done well by including the randomizing of samples, Andrews still considered it top-heavy in favor of mathematical calculations. He says that far too many books are devoted to probability, random variables, and binomial distribution. The authors confessed that “only the first nine chapters can be covered in a single semester…[which] would end the course after inference for a single mean which would leave out inference for two means and for proportions” (p. 290) as well as linear regression. The authors further suggested that the chapter on bivariant description be skipped “if student projects are not a priority”. To this, Andrews complains that if the section on design be eliminated, scatter plot and correlation would be excluded, too (p. 290).

**Essential Element**: This segment weighs the advantages and disadvantages of excising some of the heavily mathematical portions of the statistics textbook.

**Additive/Variant**: The case here is the same as that made in the segment immediately above, only that this time the practical meaning of probability and binomial distribution cannot be fully appreciated without a sense of the difference between probability and combination. Further, the “descriptive” appreciation of correlation and situations that involve “mean-proportion” requires the very mathematical forms that Andrews would rather have excised from Johnson’s second book. The mathematical formulas, calculations and descriptions make it possible for the student to *think* statistics as well as *solve* problems in statistics. The two questions are, should a revised edition include both? Or, should a revised edition exclude the mathematical portions? Since we live in such a highly technological age, where Excel and other software are now available, the answer is to do neither. Rather, there should be three texts. One text addressing statistical concepts and reasoning, should be backed up by a companion text on how to use Excel, SPSS, PSPP, etc. Also, there should be a third text that is heavily mathematical especially for those who wish to write computer programs for statistics.

**Contextualization**: Recently, I saw the movie, *Hidden Figures*, and one of the things it demonstrated to me was that the ladies who were trained in computer programming, were instrumental in writing software to solve problems in the future.

**Comment 10**:

**Quote/Paraphrase**: Andrews applauded the student-projects near the close of each of the eight chapters but found the quantity of projects rather skimpy. The first two of these pushed students “to think about sampling and experimental design”. The third prompted them to generate sample distribution by repeat sampling. The projects, in the last five of the eight chapters, are meant to push students to hurriedly specify their designs, then move on to calculations of tests and confidence intervals while disregarding exploratory data analysis, “except for the advice to ‘include a graphical display’ in one of the projects” (p. 290).

**Essential Element**: This makes the point that there is some value in applying the

tactile-kinesthetic approach to the education process.

**Additive/Variant**: As the reviewer walks us through a book on a book, which most

of us have never read, we begin to hear more and more about what should have

been pruned and what should have been spruced when it comes to student follow-

up activities based on the instruction on statistics throughout the book. If what

Andrews claims is true, the book has neglected student “reinforcement” activities

in the critical areas in parametrics, descriptive statistics and meaningfulness.

Once again, in Hidden Figures, even in the age of computer technology, all the way to artificial intelligence, the human ability to find solutions to the “novel” problems where machines fail, is undeniable. In this time when “computer pirates” are hacking into life-sustaining computer systems, if no one knows how to do the math, how can we stop the hack?

**Contextualization**: Parametric and descriptive statistics help to make statistics meaningful in a way that computer applications do not. If we are able to look at correlation in terms of a line drawn through points on a scatterplot, we should be able to give tangibility of how perfectly or imperfectly the points fit or do not fit. But once we say line, we must think of the rate of change. We are thinking of regression. And in the world where things happen, we think of them happening on an arc or a line. When the physicist talks about his inclined plane, he can make predictions about the effect that the angle of inclination (the slope) of the plane will have on the behavior of an object on it at a certain point at a certain time.

In analytic geometry we think of a line as having a specific slope (m)

and an intercept of (b). Ex. 1: y= mx + b (Line). Ex.2: m = Y1 – Y2. (Slope).

X1 – X2

In linear regression, where you will have a large number of independent and dependent variables drawn from the sample, the slope is:

m = n (∑xy) – (∑ x)(∑ y)

n (∑ x2) – (∑ x )2

**Source 4**:

Natiliia, Byshevets; Lolita, Denysova; Oksana, Shynkaruk; et al.(2019). Using the method of

mathematical statistics in sports and educational research of masters in physical

education and sport. *Journal of Physical Education and Sport, suppl*, 3(19), 1030-

1034. Doi:10. 7752/jpes.2019.s3148.

**Comment 11**:

**Quote/Paraphrase**: According to the author, the more scientific advancements there are, is the more progressive social transformation there will be in the society. As contemporary statistical activities follow their natural course through data collection and analyses, together with the identification of correlational relationships, new trends are being formed in the area of pedagogical research. Thus, the viability of hypotheses that are being presented by researchers in their attempt to find answers to problems related to the education of children is constantly being weighed and sifted in order to determine their efficacy. For that reason, the application of statistics to empirical data continues to be a vital part of teaching and performing scientific activities for the forward movement of human civilization (p. 1030).

**Essential Element**: This has to do with how applied-statistics can be utilized to help grow children K-12 into a world of rapidly advancing scientific activities in the twenty-first century.

**Additive/Variant**: We tend to see science in terms of that final product of applied human ingenuity that eventually leads to greater ease of living and cultural improvement. And so, as the author suggests, the more scientific advancements there are, the more “progressive” will the social transformation in society be. With that, little thought is given to the fact that in order for scientists to put their finishing touches on the product that comes out of their scientific activities, there is need to bring the statistician in, to test the viability and effectiveness of the product before it is adapted into a cycle of usefulness, and to make predictions about future possibilities. Hence, the scientific implements and actions around us are the results of proven hypotheses—sometimes written down, thought of, or, at other times, just spoken. So, in the end, the only important fact that exists in the teacher-student interaction over a lesson in statistics is a ***pragmatic*** one.

**Contextualization**: When I taught high school mathematics, including statistics, I followed the basic rules of pedagogy, while remaining focused on my goal to achieve a pragmatic outcome. In New York City, the forces-that-be expect the teacher to teach the entire lesson by way of the Socratic method. Therefore, with that in mind, the lesson should begin with a question and unfold through a series of hypotheses and research questions. The belief is that, if the typical Socratic kind of “question-answer” process is followed throughout the teacher’s delivery of the content, the students’ perceptions and behavior should be altered. As a result, the students will show statistical improvement on tests in the content area, over the short-term, and will at length, demonstrate specific and wholistic growth, in the long term. In the end, students must demonstrate the skills acquired, the knowledge of concepts taught, and a sense of how they are applied.

**Comment 12:**

**Quote/Paraphrase**: At the present time, a significant amount of theoretical and practical analyses of data sets has been gathered. Today, a vast cache of automated methods of statistical computations has been developed (Shelamov, 2010). In the natural, as well as in other sciences, the proper gathering of data for the conduction of scientific studies has been among the most important actions taken during research activity. But, in the areas of human centered studies, namely, in sports and pedagogical research, a large majority of the relevant data can only be gathered by questionnaire polling. In studying humans, it must be realized that the researcher, sometimes has to deal with data that may not be quantifiable. For that reason, he must realize that he might not be able to verify the level of confidence as to the objectivity of the results (Yates, 1934; McNemar Quinn, 1948, Karuskal 1952) (npn).

**Essential Element**: Because data based on human behavior and characteristics are nominal or categorical variables, they are not quantifiable, and therefore, would need to be assigned numbers within a range.

**Additive/Variant**: When we study human behavior or characteristics such as race, gender, anger, happiness, and so on, we are usually dealing with normative data. This kind of data may be assigned numbers, even though such numbers do not have quantifiable values. For example, the questionnaire could ask the respondent to enter 1 for African-American, 2 for Caucasian, and so forth. If it is a study on breast cancer research, and it is meant to screen only African-American women whose mothers were, either North Carolinian or South Texan, and numbers are assigned to all categories, a total numerical value would be required within a certain range.

**Contextualization**: I wrote a thesis for an MS degree at CCNY in the year 2000. In it, I created what I called the “Thompson College Fit” (TCF) instrument. The TCF used the Likert scale as its base. On the left, from top to bottom, there are 15 statements. The first of the statements is, “Changes in financial aid package;” the twelfth is, “Feeling academically unprepared;” and, the fifteenth is, “Not proud to be a student at this college.” The numeric choices from left to right are 100, 89, 79, 69, and 59. While, their verbal descriptions for each of the five values, in respective order, from left to right, are: “None/No,” “Seldom,” “Sometimes,” “Often,” and “Very Often.”

If student A were to pick, “None” (100) for all fifteen statements, he would

score 1500 or 100%. The inference drawn, then, would be that A’s college choice

was a perfect fit for him. On the other hand, if student B selected “very Often” for

all fifteen statements, his score would be 885 out of a possible 1500. That would be a

TCF rate of 59%. Since the New York State Education Department and the Board of

Regents have set the minimum passing rate for examinations and other activities at

65%, student B would have failed the TCF. The conclusion would, then, be that the

college student B had chosen was an absolute misfit for him.

**Comment 13**:

**Quote/**Paraphrase: “The purpose of the article is to justify the universality of using φ-test [Phi-test] criteria in scientific research dealing with physical culture and sports” (Byshavets, npn). Having accepted the universality of the Φ-test, the authors of this book call attention to the fact that this type of test exhibits some restrictions when applied to an ensemble of particles. He suggests that it is here that the F-test becomes helpful. In the questionnaire, numbers are assigned to the objects in a manner that is consistent with the rules. The data in the nominal scale are placed under their respective labels. If the numerical assignments given to each of them done in the sample under the respective labels, most experimental data can be quantified (Ashykhmina, 2009). In this scale, comparing and contrasting the different values, only allows the researcher to determine the levels at which each type of data belongs. When the F-test is brought into play, the researcher is, then, able to visualize the scaled data in terms of angles within the range of all angles within a half-turn--180o (p. 1030).

“. . . [T]he lower limits are defined in accordance with the following rules:

\*n = 2, 2n > 30;

\*nj = 3, 2n > 7;

\*n1= 4, 2n > 5;

\*for 2n > possibly any comparisons.

φ = 2 arcsine √9 [angles on 180 o are used for the variables in this article]

Computer experimental value φ-criterion by the formula.

φ null = I φ 1 - φ 2 I√( n1) (n2) where n1 , n2 –sample sizes.”

√n 1 + n2

So, an F-test is appropriately defined as angular. The first particle (P) can be

written as one of the angles within the interval 0o to 180o. if the researcher gets to where he can say “there is an effect”, and that effect can then be presented as an angle within that half-turn range in the following manner, 0 < n ≤ 180. The angles are calculated as Φ1 and Φ2 by use of the formula. Based on the rules, the initial data are plugged into the Microsoft Excel table and the experimental values of the angles were calculated by means of a formula and the particles of the samples Pi should produce an observable effect. “The analysis of literary sources showed the specialists presented methods of applying nonparametric criteria (Arkhipova, 2009). It described application of an φ-test in psychological and pedagogical researches” (Rudenko, 2012) (p.1031).

It is not surprising that “. . . nowadays, methods of test of personal data, in particular, an Φ-test, are not widely used as the basis of validity and effectiveness of results in sports and pedagogical research. The question of automating its use in the field of physical education and sport is not sufficiently covered in literary sources” (Rudenko, 2012, p. 1032).

**Essential Element:** How and when is theΦ-test used to determine the effect of one variable on another.

**Additive/Variant**: In this segment of the article, it is suggested that there is a way to use the Φ-test to uncover the “effect” that an independent variable has on a dependent variable. The implication is that the test is supposed to show the effect from the independent variable in terms of *linearity* or *curve-linearity*, if you will. My thinking is that if the line of regression is horizontal, the slope would be a “No-slope” (0/x). Such a slope would not provide a correlation coefficient that could lead to an effect that can be determined. Also, it seems that where there are multiple independent variables, the Φ-test is not adequate, so the F-test would come into play. The problem is that the “rules” provided as the means by which to perform these tests are not coherent in my estimation. There obviously needs to be more mathematical clarity in the rules and a clear algorithmic path to follow.

**Contextualization**: Throughout this article, I have looked for a clear, workable “definition” in the form of an equation or a formula, to no avail. The phi-test formula below is one example of a formula the author could have used (Wikipedia). Clearly, the formula is an “irrational expression” with a “Chain Rule” principle (syllogistic) as is shown:

φ **= . ad – bc .**

**√(a + b)(c +d)(a + c)(b + d)**

There has to be a dichotomy of variables in which at least on is nominal, such as light/dark, truth/lie, etc. Working Day broken down into night and day. Night: a =12, c = 10; Day: b =14, d = 18, So,

φ **= . (12)(18) – (14)(10) .**

**√(12)+ (14)(10)+(18)(12) + (10)(14)+(18)**

φ **= . 216 –140 . . 76 . 76 .**

**√[(26)(28)(22)(32)] = √(512512 = 715.9 = 0.10610**

Granting that the numerical values I offered here are arbitrary, the result that comes out of the phi-test at 0.106 or 11% would be a very weak positive effect. In the case of the F-test, summation of the variables could be done as indicated below (Wikipedia):

**Y = β0 + β1 X1 + β2 X2 + β3 X3 + *Ε I***

**Source 5:**

Kruschke, John K and Liddell, Torrin, M. (2018 Feb). The Bayesian new Statistics: Hypothesis

testing, estimation, meta-analysis, and power analysis from a Bayesian perspective.

*Psychonomic Bulletin & Review*; New York. 25(1), 178-206.Doi: 10. 3758/s13423-016-

1221-4.

**Comment 14**:

**Quote/Paraphrase**: These authors tell us that there is a distinct difference between the use of the traditional hypothesis-testing tools and the more novel tools that operate on the basis of “estimation” by means of **“**quantifiable uncertainty”. Today, as the winds of change continue to sweep across the landscape of scientific research, the more traditional forms of hypothesis-testing appear to be conceding their place of prominence to what are commonly spoken of as, “The New Statistics”—essentially, a form of statistical analyses that has steadily grown in popularity while avoiding the time-worn forms of hypothesis-testing (Cumming, 2014). Secondly, there is also a distinction between, what these writers refer to as the “Frequentist” and the “Bayesian techniques” of hypothesis-testing, as well as the accent placed on estimation of “confidence intervals (CI)”. This article also addresses the way the Bayesian method is used in meta-analysis to randomize control trials, and power analyses. With this, there is a shift away from null hypothesis significance testing (NHST) toward “estimation based on effect sizes, confidence intervals, and meta-analyses” (Cumming, 2014, p. 179).

**Essential Element**: This is a study of the contemporary shift away from traditional NHST approaches toward estimation, etc., by way of “The New Statistics”.

**Additive/ Variant**: At last, we are here in a highly technological age where mathematics and the magical outcomes from specialized algorithms and the proofs of theorems are king. Lately, analysists have been settling for statistical predictions resting on the viable of proofs of the “uncertainty principle”. The idea is that in order to realize the expected probability with sufficient “frequency”, the trial has to be repeated many times. For example, the expectation of getting heads when a coin is flipped is ½. Four flips could yield anywhere from ¼ heads to none at all. However, as the number of flips are increased into the hundreds, both heads and tails will tend to increase in “frequency” until the expected ½ probability begins to show up repeatedly.

**Contextualization**: If this were being taught in a high school setting, with emphasis on the teaching of the Frequentist and Bayesian techniques, it would also be necessary to teach the traditional hypothesis techniques, so that the students would be able to appreciate the points of similarity and difference. Because, the more advanced students, at least, would have already been exposed to such topics as central tendency, the teacher would be able to teach “The New Statistics” during his or her furtherance of “affective” instruction. This would also be a natural bridge to areas like standard deviation, not only from the angle of conventional mathematical, but also by way of the TI-83 or TI-84, and other computer statistical software. By those approaches, the foundation for a distribution curve, skewed, or not, would lead the way to “null hypothesis significance testing”, in the meantime.

**Comment 15**:

**Quote/Paraphrase:** The Bayesian analysis reassigns credibility in a “one-to-one correspondence” with a set of matching possibilities. Sherlock Homes appeared to have anticipated the Bayesian process in his words: “When you have anticipated the impossible, whatever remains, however improbable, must be the truth” (Doyle, 1890). To illustrate the point, he identified a string of suspects; then, after collecting new evidence, he assigned degrees of suspicion to those not eliminated by the “improbability rejection rule”. Once those who belonged to the low probability group are eliminated, the suspicion list is narrowed down to the most highly probable (p.183).

In this system, the analysis begins with “a space of possible parameter values” (Kruschke, 2018) that abound within a descriptive data model quite the way that it does in a Frequentist type analysis. The amount of belief in the parameter values, after *ignoring* the data, is known as the *prior distribution*. While the amount of belief exhibited while *accounting for* the data, is called a *posterior distribution*. “The exact re-allocation of credibility across parameter-value [*sic*] is provided by the mathematics of ‘Bayes’ rule’ ” (Kruschke, 2018, p.183).

In the frequentist approach, individuals are asked whether they agree with a

particular policy. Each respondent is treated as if he or she were being randomly

sampled to indicate that the probability of the positive response is the parameter-value being represented by the unknown variable, Q (theata). The realization is that all of these parameter-values fall within the interval of Q < 1. The analysis begins with “a space of possible parameter-values” that abound, not with the amount of belief one puts into the parameter-values while ignoring the data of the “prior distribution” (p.183). Here, the replacement for the p-value is z/N because it provides the probability that the event within the parameter has either occurred, or has not occurred. In this arrangement where we are using z/N, the p-value would not help us to come up with the “estimated magnitude of the parameter”—EMP (Kruschke, p.183).

By turning to the frequentist model, we can seek to find the “maximum likelihood estimate” (MLE). And, as such, the MLE of Q would be the actual z/N ratio. This would also be the “modal value” as well as its being pretty close to the “median value” (Kruschke, p. 183)

Essential Elements: This is a comparison between the Frequentist and Bayesian systems.

**Additive/Variant:** Routinely, when testing the null hypothesis, p < 0.05 serves as the cut-off point from which any viability of the null hypothesis is ruled out. Having collected the sample from the chosen population, without special attention to individul characteristics, randomization is assured.

The Sherlock Holmes’ analogy used by Kruschke to illustrate the Bayesian system of analysis, closely resembles the scientific system that begins by ruling out that which does not apply. This system commonly known as the “process of elimination” differs, in the Bayesian system, from what occurs in the Sherlock Holmes example by the specific order in which things are eliminated. In this system of analysis, all that is impossible is lumped together and eliminated first. At the same time, the Bayesian technique, dealing with what is first eliminated, is antithetical to the elimination process in Sherlock Holmes. For, in Bayesian analyses, like an “effect size” process, what is on the side of the actual effect being sought is first identified and then separated from the remaining results. As in the effect size approach, what remains is lumped together among the variance and other extraneous effects that fall outside of the parameter.

**Contextualization**: As a high school mathematics teacher, I have had to teach students to write mathematical proofs. Most proofs are generally “inductive” or “deductive”. In some rare situations, though, the ideal way to prove a theory or postulate, may be neither deductive nor inductive; it may be what is called an “indirect proof”. In such cases, the object is to prove that the postulate is impossible. This approach is clearly that used in Bayesian analyses.

**Comment 16**:

**Quote/Paraphrase**: In data analysis, formal mathematical models are used to describe data and their parameters. For instance, a shower head could be tilted in a manner that causes the water (data) to be sprayed out at a certain “angle of inclination”. We could treat a particular angled-spray and the type of spray-pattern created by the setting of the nozzle as the two parameters. So, the data observed are manipulated to fit the appropriate mathematical machine to which its parameters are set. In this case, the object is to come up with other settings of the parameters that simulate the empirical data **(p**.179).

Likewise, if we decided to measure the intelligence quotients (IQ) of a group of people, we could make a histogram of the scores—one that has a single mode and is symmetric—one that could be translated into a “bell curve”. The mean and the scale, or the standard deviation, would, thus, be the parameters. By comparing these two parameters with the angling of the spray from the nozzle, and by setting the type of spray-action set by the control knobs, we are able to obtain a sense of how these IQ scores could be similarly manipulated to work in a mathematical machine that is aptly suited for generating normally distributed data (p. 179).

**Essential Element**: This has to do with Bayesian analysis in real world situation.

**Additive/Variant**: This situation with the shower head must first be visualized from a purely geometrical standpoint. As such, a number of geometrical adjustments can be made that simulates the angled spray that was originally observed. The simulations could be one in which angles and spray adjustments could be set randomly to suit the height and comfort of the person using the shower. Or it could be set in random similarity by “direct variation”. This latter adjustments of the angle and spray would be more practical for people of different sizes and comfort levels, and a better suited comparison to the IQ parameters of mean and standard deviation. In this respect, we are using the word “similarity” in a non-mathematical as well as a mathematical way.

**Contextualization**: If the original angling of the shower head were at a forty-five degree angle and the spray were set at a dilation of 10 inch-diameter at floor- contact, they could be adjusted to a 40o angle and a 8.9 inch-dilation. This is a mean-proportion or direct variation (x1: y1 = x2 :y2) relationship that makes it perfectly “similar” to the original 45o : 10 ratio. Otherwise, the angle and spray-dilation parameters could be disproportionately paired as they would naturally occur in a real-life situation. When we compare this to IQ scores of a random sample, of let us say 20 people in a particular population, you would expect the sample number to range probably from 90 to 140. That would make it possible to find central tendency and standard deviation of frequent repeats.

**Comment 17**:

**Quote/Paraphrase**: Individuals are asked whether they agree with a particular

policy. Each respondent is treated as if he or she is being randomly sampled to

indicate the probability of the positive (“Yes”) response as the parameter value

being represented by the unknown variable, Q (theta). The idea is that all of these

parameter-values fall within the interval, 0 < Q < 1. Using the Bayesian model, the

goal is to re-allocate credibility to the values in the previously mentioned interval except for those values that are not consistent with the data.

To begin with, values of Q are harvested through the process of “prior

distribution”, so that the number of positive responses (z) are pulled out of

the total (N) number of responses. Credibility, is, then, reallocated to the “Yeses”

(z/N) that coincide with what is actually observed in the data. At this, it so happens,

that some of the small values of *theta* are very inconsistent with a good many of the

positive responses (the “Yeses”) in the data. And so, the small values of theta should wound up being assigned to the “low posterior credibility (LPC)” bunch.

Once the actual data is accounted for, the “posterior distribution” begins to become acceptable and the level of credibility is elevated. As the posterior distribution spreads out, in accordance with the standard deviation, the “uncertainty” of the small parameter values becomes more evident. Since the highest density interval (HDI) bunches up within 0 < Q < 1, it contains parameter-values of higher “probability density”. The probability outside of Q < 1, including those within 0 < Q progress from low to very low, and therefore, they fall within the realm of “uncertainty” which makes them very likely to be “improbable” (pp. 182-’84).

**Essential Element**: How the Bayesian analysis works by the elimination of the

improbable.

**Additive/Variant**: Simply put, Bayesian analysis comes close to a purer form of inferential logic than Frequentist analysis does. The question of what the Bayesian process is all about is explored. Bayesian analysis is a method of hypothesis testing done on the front side by the “process of elimination”. By turning to the frequentist model, we seek to find the “maximum likelihood estimate” (MLE). And, as such, the MLE of Q would be the actual z/N ratio. If 14 out of a group of 18 participants provided the expected answer, the ratio would be, 14/18 = 0.77% for the MLE. This would also be the “modal value” as well as it being pretty close to the “median value”. In frequentist methodology, we represent the “uncertainty” of the estimated parameter-value by first coming up with a *confidence interval* (CI). If we make the confidence interval 95%, all the parameter values within it would be safe from rejection, but those values outside are open to use of the posterior distribution, the uncertainty of the estimate should be evident in the breadth of its spread. However, at one standard deviation to the right of the population mean (µ), any assumed certainty of the p-value indicated by Q < 1 on the outside right, toward the upper bound would be rejected. Thirty-four percent on the left of the mean, at approximately 0.05, would render all values moving toward 0.0 to be highly uncertain. To fully illustrate this, it would be necessary to present it visually in the form of a normal distribution.

**Contextualization**: Twenty teachers (N) of a high school senior-class are asked whether the entire class will satisfy the graduation requirement this year. Sixteen of them (z) responded in the affirmative. The proportion of those who responded positively is z/N, which gives us an actual numerical result of 16/20 or 80% probability. So, while the Greek letter, Q (theta) is used to represent the probability of the positive responses, and the Q**null** represents those answers that are not “Yeses”, there is only certainty that the remaining 20% probability applies, not to the “Nos,” but to both the “No s” as well as the “Not sures”.

**Source 6**:

Fernando, Jason. (2023, Oct. 18). The correlation coefficient: What it is, what it tells investors.

*Investopedia* (Reviewed by Margaret James and Fact checked by Vikki Velasquez).

**Comment 18**:

Quote/Paraphrase: According to Jason Fernando (2023), whatever standard scale

has been set based on the extent of Pearson r’s ability to tell the level of strength in

the relationship between the two variables, it is not set in stone. But rather, it is

set upon a numerical scale which indicates the strength of the relationship between

independent and dependent variables in a manner that differs from one family of

disciplines to another. He argues that the strength of the correlation coefficient that

is considered significant in physics and chemistry with a threshold on both sides of

zero where (0.94 r) (r -0.94) in order for it to be considered meaningful. At

the same time, he claims that in the social sciences, a correlation coefficient of -0.5

or +0.5 is considered moderately significant (Fernando, 2023, p.n.).

**Essential Element:** What is the lowest correlation coefficient that is

meaningful in research?

**Additive/Variant:** Based on Fernando’s explanation of the function of

the correlation coefficient, Pearson r should indicate what the “strength and

direction” of the relationship between the independent and dependent variables, in

a hypothesis, is, in respect to the near collinearity of the datapoints with the “line of

best fit” or the line of regression.

**Contextualization**: According to Fernando, a correlation coefficient of 0.8

or 80% would not even be considered in chemical and physical science research,

while a correlation coefficient of 0.5 (50%) would be perfectly fine in social science

research. Although education, in the broad sense, includes both chemistry and

physics, education research belongs to the category of social science. By that fact,

then, a correlation coefficient of 0.5 would definitely be applicable in educational

research.

**Comment 19**:

**Quote/Paraphrase**: Although a “correlation coefficient”, by itself, is not an indication that the behavior of one variable affects the behavior of the another, the dependent variable (y) is frequently spoken of in mathematics as a “function” of the

independent variable (x). So, in mathematics, at least, the behavior of x affects the behavior of y. What the correlation coefficient is supposed to help us see is

whether or not we can make a reasonable assumption about the strength of the

relationship between the x and the y variables. Also, the convention that sets up

how Pearson r helps the researcher to determine the strength of the relationship

between the two variables, differs among the many families of disciplines

(Fernando, 2023, p.n.).

Fernando (2023) speaks of a process he refers to as the “Coefficient of

Determination” which he describes as “R-squared”. He writes that the process shows

“. . .what portion of the variation in the dependent variable is attributableto the

independent variable” (Fernando).

**Essential Element**: What is the “Coefficient of Determination” and what does it do in

statistical research?

**Additive/Variant**: In effect, the “Coefficient of Determination” is R2. By using the

Coefficient of Determination, the researcher is able to identify the amount of effect

that the independent variable causes to the dependent variable and what portion of

the behavior of the dependent variable is caused by variances and unspecified

variables.

**Contextualization**: Jason Fernando tells us that to find the “Coefficient of

Determination” one has to, first, figure out what the Pearson r is and what the

square of it turns out to be. For instance, if r = .3, the “Coefficient of Determination”

would be equal to .09. We recognize also that the “absolute value” of Pearson r

reaches its “limit” at 1. What this means is that any possible value of r that could

exist is confined to the interval of the inequality -1 r 1. When the correlation

coefficient is r =.3 and the square of r is .09, R2 can be checked by simply finding the

square root of R2. Therefore, we could actually do a check on R2 = .09 by simply

applying the inverse operation of .32 as follows, .09 = .3.

**Source 7**:

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillside, N.J.,

Lawrence Erlbaum Associates.

**Comment 20**:

Quote/Paraphrase: Cohen (1988) writes that, presently, “Effect Size” has been

gaining popularity as the go-to system for identifying the effects from variances

apart from the possible effect that the independent variable may have on the

dependent variable. Previously, the popular approach in research testing had long

stressed the utilization of test-statistics to isolate and weigh such effect.

But now, researchers, especially those working in medical science research, are

tending to utilize the process of “Effect Size” more and more. In Cohen’s description

of “Effect Size”, he says that it is (Eta-squared) which he claims is “. . . analogous

to r2 but is used for t-tests and ANOVA. Like r2, it reflects independent variable”

(Cohen, 1988).

**Essential Element**: This addresses the workings of “Effect Size” in statistic testing.

**Additive/Variant**: While Jose Fernando (2023) refers to the process he calls

“Coefficient of Determination,” and identifies it in mathematical terms as R-squared,

he writes that the process shows “. . .what portion of the variation in the dependent

variable is attributableto the independent variable” (Fernando). What Fernando

refers to as “Coefficient of Determination” is exactly what Cohen and others call

“Effect Size”.

**Contextualization**: In an effort to find the strength of the “relationship” between a

certain pair of independent (x) and dependent (y) variables, we plug the data

into Excel’s “Correlation” function and obtain a Pearson r of 0.6457. We recognize

that, while this correlation coefficient, by itself, does not tell us anything about the

“effect” that the x-variable may have on the y-variable, we recognize that, having

found the value of r, we are able to take the next step. We are able to do the math,

and by so doing, determine what effect, if any, does x have on y, and how much of

y’s behavior is determined by errors or something else, outside of x. As a result, the

Effect Size process is applied. With r = 0.6457, R2 is 0.41692849. This tells us that

41.69% of the behavior of y is a result of the effect from x. By this fact, we are able

to determine that 58.31% of the remaining effect on y is the result of variances,

which may include errors and other intervening variables.

**Source 8**:

Reichard, Joshua (2020). *Statistics as a language: Overview of statistical tests and hypothesis*

*testing* (tutorial 2). [Video].YouTube. [Time: 0.03/21:36].

**Comment 21**:

**Quote/Paraphrase**: “. . . .[I]f our sample data actually represented the population

for a given hypothesis, or if they only represent the sampling errors. . .given we are

conducting the research, the prevailing assumption is that our hypothesis about the

population is false—no one has demonstrated otherwise. So, it is up to us to reject

the null hypothesis, not ‘prove’ the alternative hypothesis. Remember the burden of

proof is on us!” (Reichard, 2020).

**Essential Element**: This has to do with the researcher’s burden to prove that his

assumption is true by using the null hypothesis to prove that what it is not is false.

**Additive/Variable**: It does appear that this passage is pointing to a convention that

binds scholarly researchers to a condition of honesty that might not otherwise have

existed if there were not the imperative that the research begin with a null

hypothesis (H0). In Reichard’s suggestion, the researcher has the responsibility to

reckon with the probability that his hypothesis (Ha) could be false, before moving

forward. And so, until we test to see whether it is true that the opposite of our

hypothesis about the population we wish to study is false, we do not have the right

to proceed. The question is should one expect to see the testing of the null

hypothesis in all quantitative dissertation?

**Contextualization**: This use of the null hypothesis in scholarly research closely one of

the three basic forms of mathematical proofs. The three most popular types of

mathematical proofs are, deductive, inductive, and indirect. The latter of the three

types is the one least frequently taught in high school mathematics classes around

the country, but it is precisely what researchers do when they test the null

hypothesis. Simply put, an inductive proof is one in which we prove what a thing is

not in order to prove that it is.

**Comment 22**:

**Quote/Paraphrase**: While there is a number of different types of errors in statistics,

two very common errors in hypothesis testing are the Type I and Type II errors. The

Type 1 Error occurs when a null hypothesis that should be rejected, is not. This is

typically known as a “false positive”. The Type II Error is the exact antithesis to the

Type I Error, which is commonly referred to as a “false negative” (Reichard, 2020).

**Essential Element**: This is a comparison between Type I and Type II errors with

examples.

**Additive/Variant**: Some researchers tend to test their null hypothesis before

anything else. At times, they run the test even prior to sample collection. In such

cases, he usually chooses what the p-value is going to be, in advance. But

there is a downside to using this approach. Researchers who use this approach

are more likely to commit Type I or Type II errors. Type I and II Errors occur

when there is a high level of alpha and beta probability, respectively. Whenever a

hypothesis test fails to reject the null hypothesis, even though the probability ()

says it should, we have a Type I Error. On the other hand, Type II Error is the failure

of the hypothesis test to verify that the null hypothesis is false, even though, the

probability of Beta () indicates that the alternate hypothesis is true.

**Contextualization**: One of the causes of the Type I Error is that it often occurs when

there is confusion between “statistical significance” and “practical significance.” For

example, a tennis coach is concerned whether his young thirty-man, Olympic-level,

tennis team should be coached to play the backhand crosscourt topspin, single-

handed or double-handed. The coach organized the top twenty members of the team

into two equal groups of ten, group A and group B. The tennis players in Group A,

going forward, would be coached to hit the double-handed, crosscourt, topspin

backhand. While those in group B, were being coached to hit the single-handed,

crosscourt, topspin backhand. A dozen best of two, three-setters were played head-

to-head within each group, then, randomly played between players across the two

groups. After many repeats of the test, it was found that the double-handed

backhand consistently had 5 to 6 more winning shots (“winners”) than the single-

handed, backhand.

It was found that 75% of the players in Group A reported a peculiar hip injury,

while 100% of Group B reported no such injury. Yet, prompted by the outcome of

the doublehanded winners over those of the single-handed ones, the coach decided

to switch the entire team to hitting the double-handed crosscourt topspin backhand.

The alternate hypothesis was, "There is statistically significant difference between

the double-handed, crosscourt, topspin backhand and the single-handed, crosscourt,

topspin backhand.”

For, when the reconstituted team competed against other outside opponents,

the team’s win/loss ratio turned out to be 2/8 as against a 7/3 win/loss ratio,

historically. The coach, having switched to the double backhand stroke, did not take

into consideration the “practical significance” of the fact that there was a spate of

sit-outs as a result of hip injuries. This is a Type I Error.

From the effect of style of play on winning tennis, an alternate hypothesis (Ha)

created, states that, “There is statistically significant difference between hitting a

double-handed, crosscourt, topspin backhand and hitting a single-handed,

crosscourt, topspin backhand.” Should the null hypothesis (H0) be rejected, the

alternate hypothesis would automatically be true based on the tests conducted

above, between Groups A and B. If, in conducting the hypothesis-test to make that

determination, the null hypothesis would fail to be “rejected”, and the alternate

hypothesis would be false.

But what if such a failure to reject H0 is, itself, error? In this case, the coach’s

decision about the direction the team should take could be disastrous. The coach

could erroneously decide that no one of the two styles of play should be used as the

staple for making the team perform at its highest level.

The coach may take the extreme position of concluding that even though the

larger number of winners were made by the double handed, crosscourt, topspin

backhand, the hip injuries among this group had off-set the benefit of playing this

style of tennis. Although the null hypothesis was not rejected, should have been in

light of the hip injuries and the team’s subsequent failure in competitive. Since this

decision would have been made because of a false negative, it would be, in effect, be

a Type II Error.

**Source 9**:

Reichard, Joshua (2020, March 8). *Statistics as a language: Introduction to statistics*.(Tutorial

1). [Video].YouTube.https://www.youtube.com/watch?v =FGeUc2F17ek [Time =

29:20].

**Comment 23**:

**Quote/Paraphrase**: According to Reichard (2014), “Statisticians use the term,

‘distribution,’ to refer to how data will look. They are graphed in what is called a

histogram. The y-axis represents the frequency or number of people who have a

particular salary, and the x-axis, going horizontally shows the monthly salary for each

group of people. So, in this example, it looks like about 25 people make just over

$800 per month and very very [sic] few people make $1500 per month . . . . [T]he

most of the group make just over $900 per month . . . .” (Reichard, 2014; video).

**Essential Element**:

**Additive/Variant**: Contrary to the above view, in statistics, the term “distribution”

has more to do with how data are spread out, or literally, how they are distributed,

than it has to do with how they are presented graphically. When presented

graphically, in the form of a histogram, the data is spread out from the “origin (0,0)”

with a precise “periodicity” on the positive side of the horizontal, or x-axis. Out of

each of these periods, whatever is occurring, rises up to a y-value that indicates how

frequently the event in that particular period is occurring. In alignment with

Reichard’s example, the period on the x-axis, to which $800 per month is assigned,

rises to the number 25, on the y-axis, which indicates how frequently that periodic

event on the x-axis occurs; that is that as many as 25 people make $800 per month.

So, the pattern of varying frequencies of the events occurring on the x-axis is

indicated by values on the y-axis.

**Contextualization**: To get a clear sense of the relationship between the Independent variable (x) and the dependent variable (y), could probably be best achieved through mathematics. If we reduce the statistical relationship between independent and dependent variables to a single point in Cartesian plane, we could say that every point consists of an abscissa (x) and an ordinate (y), where order is important. What that means is that if a point is identified as (2, 4), by convention, x = 2 and y = 4; never the other way around. By that convention, then, we may say that y is the function of x, but never x is the function of y.

**Comment 24**

**Quote/Paraphrase**: In the linear equation, “y = 2x -7,” when x = 2, what results is y =

-3. If we were to restate the linear equation (y = 2x -7) as a function, we would say,

f(x) = 2x- 7. In this case, the point (2, y) would become (2,-7) which indicates that the

y-variable exists only because of the behavior of the x-variable under the given

condition. Therefore, in the Reichard-example, the group of 25 people results solely

from its relationship to the salary of $800 per month, indicating that y is a function of

x.

**Comment 25**:

**Quote/Précis**: Reichard (2014) states that by the time the histogram is completed,

the researcher should have acquired a visual sense of whether this particular

histogram can be translated into a “normal distribution”. To do such a translation,

the sample or population “mean”, whichever is applicable, together with the

“standard deviation”, must first be determined. The horizontal axis should be

treated as “the number line”, with the “mean” occupying the center, at zero, while,

the standard deviation spreads the data out, over a range, at even periodicities.

The highest frequency from the histogram (the “mode”), is then arranged in a

manner that allows it to rise out of the “mean”, up to the apex of the graph, while,

each of the lower frequencies falls progressively down to its respective period, on

the left and the right sides of the mean to fit into the bell-shaped sketch around the

histogram. If this arrangement of the data, in the histogram, takes the rough shape

of a symmetric bell curve, we have a “normal distribution”.

What this means is that a chunk of data, in the middle of the bel-curve, from

from top to bottom, is spread out at one standard deviation, on either side of the

mean (-34% and 34%) . This entire spread of sample data across the mean, maps out

the median at a total of 68%. Should the data continue to spread in the two

opposite direction away from the center (the mean), if, or until they reach the

second standard deviations, at -13.5% and 13.5%, the total spread of the sample

data would be 95%. In such a case, the *p-value*, or the difference between the

applicable sample data at play (95% of them) and theprobable certainty (100%)

would leave a result of 5%. Through Reichard’s graphic translation of the histogram

into the bell curve, we are able to see that, since the remaining 5% of the data has

to be divided between two tails, on either side of the curve, the value in one tail

would have to be less than 0.05. In this situation, the researcher would be forced to

reject the null hypothesis in favor of the alternative hypothesis.

**Essential Element**: This has to do how one finds the p-value in the course of

translating a histogram into a bell curve.

**Additive/Variant**: I would think that even though this segment of the video (Tutorial

is very well explained, it would have been helpful if the tutor had used a cursor

or an arrow as a pointer. Also, something is lost in the translation. what is lost is that

the tutor never explained how the frequencies are found on the left and the right

sides of the mean once the translation is done.

**Contextualization**: Although I taught high school statistics only from the point of view

of pure, rather than, applied mathematics, I found that in teaching students how to

translate histograms into normal distribution curves, it was necessary to teach them

how to find the frequencies on the left and, on the right of the mean. I taught them

that once the histogram is translated into a bell curve, the number of the standard

deviation on the left should be subtracted from the mean and added to the mean on

the right, in order to find the frequencies.

For example, the first period to the left of the mean is ( the second is

( is(), and so on. While on the right side, it is (

SD = 6. \_\_\_\_\_\_.\_\_\_\_\_.\_\_\_\_\_\_.\_\_\_\_.\_\_\_\_\_.\_\_\_\_\_.\_\_\_\_\_.\_\_\_\_\_\_\_\_\_\_

Mean= 20 2 8 14 20 26 32 38

In his effort to explain why a hypothesis test is necessity in scholarly

research, Reichard goes on to illustrate that, while a statistical study of the entire

population would be ideal, such a study would be impossible. So, the next best thing

would be for us to randomly select a sample that would very probability support our

assumption about the population, as a whole. To that end, having collected such a

sample, we would proceed to formulate and test a null hypothesis (H0) that is the

exact anthesis of the alternative hypothesis we wish to threat as the basis of our

study. Therefore, with the test-result from the null hypothesis at p < 0.05, we should

feel a sufficiently high “level of confident” that the 95% probability remaining would

be the green light for us to proceed with our sample as being very closely

representative of the population.

Works Cited

Andrews, Douglas M. (1999 Aug). The American statistician. *Journal of Statistics*

*Education.* Alexandria. 53(3), 291.

Bala, Jyoti. (2016 Nov). Contribution of SPSS in social sciences research. *International Journal of*

*Advance Research in Computer Science*. 7(6) (Special Issue).

[www.ijjarcs.inform](http://www.ijjarcs.inform).

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillside, N.J.,

Lawrence Erlbaum Associates.

Fernando, Jason. (2023, Oct. 18). The correlation coefficient: What it is, what it tells investors.

*Investopedia*. Reviewed by Margaret James and Fact checked by Vikki Velasquez.

Hardwickes, Tom E; Goodman, Steven N. (2020, Oct). PloS One; *San Francisco*. Vol. 15(10).

e239598. doi: 10: 1371/journal.pone. 0239598.

Kalaian, Sema A; Kasim, Rafa M. (2014 Mar). A meta-analytic review of studies of the

effectiveness of small-group learning methods on statistics achievement. Journal of Statistics Education. Alexandria 2(1), 1-19.

Doi: 10. 1080/10691898.2014.11889691.

Kruschke, John K; Liddell, Torrin M. (2018 Feb). The Bayesian new Statistics: Hypothesis testing,

estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review*; New York. 25(1), 178-206.

Doi: 10. 3758/s13423-016-1221-4

Natiliia, Byshevets; Lolita, Denysova; Oksana, Shynkaruk; et al.(2019 May). Using the method of

mathematical statistics in sports and educational research of masters in physical education and sport. *Journal of Physical Education and Sport, suppl*. 3(19), 1030-1034. doi:10. 7752/jpes.2019.s3148.

Reichard, Joshua (2020, March 8). *Statistics as a language: Introduction to statistics*.(Tutorial

1). [Video].YouTube. https://www.youtube.com/watch?v =FGeUc2F17ek [Time =

29:20].

Reichard, Joshua (2020). *Statistics as a language: Overview of statistical tests and hypothesis*

*testing* (tutorial 2). [Video].YouTube. [Time: 0.03/21:36].