Journal of Strategic Security

Volume 2
Number 1 *Volume 2, No. 1: February 2009*

Article 3

Teaching Bayesian Statistics To Intelligence Analysts: Lessons Learned

Kristan J. Wheaton
Mercyhurst College

Jennifer Lee
Mercyhurst College

Hemangini Deshmukh
Mercyhurst College

pp. 39-58

Teaching Bayesian Statistics To Intelligence Analysts: Lessons Learned

Author Biography

Kristan J. Wheaton, Department of Intelligence Studies, Mercyhurst College

Jennifer Lee, Department of Intelligence Studies, Mercyhurst College

Hemangini Deshmukh, Department of Mathematics, Mercyhurst College

Abstract

The Community must develop and integrate into regular use new toolsthat can assist analysts in filtering and correlating the vast quantities of information that threaten to overwhelm the analytic process…—Commission on the Intelligence Capabilities of the United StatesRegarding Weapons of Mass Destruction (The WMD Report)1Unlike the other social sciences and, particularly, the physical sciences, where scientists get to choose the questions they wish to answer and experiments are carefully designed to confirm or negate hypotheses, intelligence analysis requires analysts to deal with thedemands of decision makers and estimate the intentions of foreign actors, criminals or business competitors in an environment filled with uncertainty and even deliberate

Teaching Bayesian Statistics To Intelligence Analysts: Lessons Learned

Kristan J. Wheaton
Jennifer Lee
Hemangini Deshmukh
Mercyhurst College

The Community must develop and integrate into regular use new tools that can assist analysts in filtering and correlating the vast quantities of information that threaten to overwhelm the analytic process...
—*Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction (The WMD Report)*[1]

From the 9/11 Commission Report to the WMD Report to the Intelligence Reform and Terrorism Prevention Act of 2004, there has been a renewed emphasis on an improvement in intelligence analysis and the tools and methods that analysts use to create their estimates. Congress, frustrated with recent intelligence failures that were linked, at least partially, to poor analysis, decided to legislate certain aspects of the analytic process by mandating the use of a method commonly called "Red Teaming" in the 2004 Act:

> Not later than 180 days after the effective date of this Act, the Director of National Intelligence shall establish a process and assign an individual or entity the responsibility for ensuring that, as appropriate, elements of the intelligence community conduct alternative analysis (commonly referred to as "red-team analysis") of the information and conclusions in intelligence products.
> —*Intelligence Reform and Terrorism Prevention Act of 2004*[2]

This and other methods used routinely by intelligence analysts suffer from a number of problems. Unlike the other social sciences and, particularly, the physical sciences, where scientists get to choose the questions they wish to answer and experiments are carefully designed to confirm or negate hypotheses, intelligence analysis requires analysts to deal with the demands of decision makers and estimate the intentions of foreign actors, criminals or business competitors in an environment filled with uncertainty and even deliberate deception. For these reasons, those unfamiliar with the challenges of this discipline often criticize intelligence analysis methods:

Produced by The Berkeley Electronic Press, 2009

> It's time to require national security analysts to assign numerical probabilities to their professional estimates and assessments as both a matter of rigor and of record. Policymakers can't weigh the risks associated with their decisions if they can't see how confident analysts are in the evidence and conclusions used to justify those decisions. The notion of imposing intelligence accountability without intelligent counting—without numbers—is a fool's errand.
>
> *—Michael Schrage, Senior Advisor To MIT's Security Studies Program*[3]

Scientists base their criticisms on a deep understanding of the power of statistics. Traditional statistics—the kind of statistics that one commonly studies in undergraduate or graduate programs—is based largely on normal distributions, structured data sets, linear regression analysis, null hypothesis testing and the like. It produces stunning results and is responsible for many of the advances in the hard and soft sciences.

This type of analysis, on the other hand, is normally considered inappropriate for intelligence analysis. The data collected for intelligence analysis are largely unstructured, often incomplete or deceptive, and rarely capable of interpretation by these scientifically acceptable methods. Even the most quantitatively oriented analysts acknowledge that there are some, perhaps many, problems that do not lend themselves to numerical analysis.[4] One response to this issue is to claim that intelligence analysis is mere guesswork, unscientific and prone to massive failures.

Another response is to look for alternative structured methods, methods that have a basis in science but allow for the uncertainty of the traditional intelligence data set. One of the methods often recommended to analysts by the scientific community is to apply the statistical theory of the 19th century clergyman and mathematician, Thomas Bayes. Schrage recommends Bayes Theorem and calls it a "powerful tool to weigh evidence."[5] Bruce Blair of the Center For Defense Information also advocates Bayes to the intelligence community and calls it a "rigorous approach."[6]

Bayes, a non-conformist Minister and a Fellow of the Royal Society, is largely remembered today for his work on non-traditional statistical problems.[7] Specifically, the Bayesian Method depends "on taking some expression of your beliefs about an unknown quantity before the data was available and modifying them [the beliefs] in light of the [new] data."[8] Such an approach appears to have immediate applicability to intelligence problems. Previous analyses, or an analyst's own experience, often pro-

http://scholarcommons.usf.edu/jss/vol2/iss1/3
DOI: http://dx.doi.org/10.5038/1944-0472.2.1.3

vide an intuitive idea of the range of probabilities inherent in a particular problem. While Bayes has its detractors, the method provides the math that would allow analysts to update their current beliefs in a logically consistent and scientifically defensible way.

The difficulty, as always, lies in the details.

People do not appear to be natural Bayesians, i.e. they do not seem to follow Bayesian reasoning when making a decision.[9] Furthermore, Bayes seems difficult to teach. It is generally considered to be "advanced" statistics and, given the problem that many people (including intelligence analysts) have with traditional elementary probabilistic and statistical techniques, such a solution seems to require expertise not currently resident in the intelligence community or available only through expensive software solutions.

The intelligence community has toyed with Bayes before. In 1978, Nicolas Schweitzer reported on an experiment using Bayesian analysis.[10] At the time, according to Schweitzer, the technique was already used extensively in imagery analysis and was thought (by no less than the then Director of Central Intelligence, William Colby) to have potential in political analysis. The experiment, in fact, confirmed the utility of Bayes in less technical analysis and despite several misgivings, including, among others, problems with data, problems over time and the potential for manipulation, Schweitzer pronounced the technique "a useful adjunct to traditional analysis."[11] Despite this endorsement, Schrage's modern critique tracks well with one of the author's own experience: Most analysts have never heard of Thomas Bayes, much less possess the ability to apply his theories to non-technical analytic problems.

This paper reports on efforts to change that. Can Bayes be simplified such that it can be taught? More importantly, can Bayes be taught in such a way that entry-level analysts can use the method at any time and under any circumstances? This paper outlines a series of studies and experiments over the last year that try to do just that and, based on those experiments, suggests new lines of research that might well provide the answers to these problems.

Addressing the needs of junior analysts has never been as important as it is now. Due primarily to a hiring freeze in the 1990's, analysts inside the intelligence community tend to be either quite senior (and near retirement) or quite junior (hired within the last five years). For example, if Congress allows the FBI to hire all of the analysts it needs to replace those retiring in 2007, approximately one third of its analysts would have less than two years experience.[12] If Bayes is a potential solution, then figuring out how to teach it to young analysts is a critical concern.

A Brief Introduction To Bayes

Formally, Bayes Theorem[13] looks like this:

$$P(H|D) = \frac{P(H)*P(D|H)}{(P(H)*P(D|H))+(P(NH)*P(D|NH))}$$

Where:

P(H|D) is the probability that the hypothesis is true given the data.

P(H) is the probability of the hypothesis being true.

P(D|H) is the probability of the data given that the hypothesis is true (hit rate).

P(NH) is the probability that the hypothesis is not true.

P(D|NH) is the probability of the data given that the hypothesis is not true (false positive).

For too many analysts this is inaccessible gobbledy-gook. It is, however, a fairly straightforward equation. Bayes Formula is based on conditional probabilities, that is the probability of an event happening given the probability of another event happening. For example, take a coin toss. The probability that a fair coin will land heads up is 50%, therefore making the chance that the coin will land tails up is also 50%. These two probabilities are the initial probabilities (which are P(H) and P(NH) in the formula above). We also know that no matter how many times we toss the coin, the chance that the coin will land heads up will always be 50%, no matter if the coin landed heads up in the previous toss or not. This makes P(D|H) and P(D|NH) both equal to 50%. Since the probability of a coin either being heads up or tails up does not change no matter how the coin landed previously, no updating is needed, therefore indicating that the conditional probabilities will not change. Bayes Formula is able to take this into

http://scholarcommons.usf.edu/jss/vol2/iss1/3
DOI: http://dx.doi.org/10.5038/1944-0472.2.1.3

consideration and shows that the probability of a coin landing heads up will be the same as the initial probability of that coin landing heads up:

$$P(H|D) = \frac{(.5 * .5)}{((.5 * .5) + (.5 * .5))}$$

$$P(H|D) = \frac{.25}{.25+.25}$$

$$.50 = \frac{.25}{.50}$$

However, when the probabilities are not equal, Bayes Formula shows how the probability of an event happening can be affected by the probability of another event occurring.

In terms of intelligence, it starts with the quite normal idea that all analysts have a sense of the odds regarding their areas of responsibility. For example, if you asked an area expert what were the odds that two countries in his or her area of expertise would go to war in the next two years, that analyst could probably give you a pretty good ballpark estimate. (For example: "There is only a 10% chance that they would go to war.")

What happens, however, when a new piece of information comes in concerning the two countries? All analysts, but particularly young or entry-level analysts, run the risk of giving in to a number of possible cognitive biases. Two of the most common, for example, are *vividness* and *recency* biases. Both of these biases cause humans to give greater weight to vivid or recent events. Bayes, however, allows the analyst to rationally update his or her previous assessment in light of the new evidence by stating that the new probability that a particular hypothesis is correct (given new evidence) is equal to the old probability multiplied by the odds that the new event was caused by that particular hypothesis. For example, imagine that an analyst receives a piece of information he or she assesses as having a 90% likelihood of being accurate. The piece of evidence states that war between the two countries is now eminent, although the analyst had previously thought there was only a 10% chance. A non-Bayesian analyst might fall prey to a number of biases and over-react; however, a Bayesian analyst would update the probability of war using Bayes Theorem and realize that even with the new information, the real chance of war between the two countries is only 50–50.

While this may seem counterintuitive at first, as Schweitzer noted in 1978, "much of the acceptance of our current efforts is due to earlier experiments applying Bayes to the analysis of historical intelligence situations."[14] Bayes works, and for traditional intuitive style analysts, this is bad news. It gets worse, however.

Consider another example: this time, the classic example of the application of Bayes, medical testing. Imagine a disease that exists in only 2% of the population and a test that has an 80% accuracy rate. Given a positive test, what is the likelihood that you actually have the disease? Only about 7.5%. If this seems wrong to you, you are not alone. Many people, doctors included, allow biases to affect their judgment and would likely guess that receiving a positive test would indicate that the person had a high probability of having the disease, most likely very close to 80%. This is because people are influenced by the high accuracy rate of the test and do not take into consideration that only 2% of the population has the specific disease and that 20% of the time, the test incorrectly identifies someone as being positively diagnosed. In other words, where base rates are relatively low, it is easy for the number of false positives to overwhelm the number of true positives. When these base rates and false positives are taken into consideration, as is done in Bayes Formula, the answer is more accurate.

Natural Frequencies

Most people do not see how the above example works. Health professionals, from doctors to counselors, routinely make incorrect assessments of probability when confronted with these types of problems. Moreover, the bulk of previous research (as cited earlier) suggests that there is little that anyone can do about the average human's inability to take the base rate of occurrence into account when assessing probabilities. Bayes is cognitively challenging and even disconcerting; therefore, a user-friendly way that helps people understand would be extremely useful (in much the same that Sudoku is capable of getting people to perform matrix algebra by having them play a game that is easily understood and fun).

Enter Gigenrenzer. Gerd Gigenrenzer, a professor of psychology at the Free University of Berlin, along with a number of collaborators, hypothesized that the reason humans did so poorly on tests of their Bayesian reasoning skills was that the problems had been expressed in terms of traditional statistics.[15] He believed that traditional statistical thinking was

http://scholarcommons.usf.edu/jss/vol2/iss1/3
DOI: http://dx.doi.org/10.5038/1944-0472.2.1.3

a skill that only some humans acquired and thus was not natural for most humans. He extended this logic one further step: What if you use natural frequencies (e.g. "5 times out of 100") instead of traditional statistical formulations (e.g. 5%) in framing the problem? Will this help people be better Bayesian reasoners?

Psychologists have known for years that "framing" is a powerful tool for eliciting predictable responses from individuals. Particularly egregious examples pop up during every election, of course, but even innocuous and mathematically equivalent statements yield startlingly different results when framed as a loss or a gain.[16] Beyond mere framing effects, however, Gigenrenzer and his collaborators posed a hypothesis founded in the ecological thesis that posits that if the mind has evolved to reason in the Bayesian manner, it is likely "tuned" to natural frequencies.[17] Gigenrenzer sought to use these phenomena to overcome the perceived inability of humans to use Bayesian reasoning.

Gigenrenzer and his collaborator, Ulrich Hoffrage, tested two groups of doctors. In the first case, Gigenrenzer presented a problem (similar to the one described above) to one group of doctors using traditional statistical expressions. The doctors did quite poorly and these results tracked with numerous earlier studies of Bayesian reasoning. In the second group, Gigenrenzer used natural frequencies—and the results were startling. Framing the question with natural frequencies led doctors to score significantly better on the test of their Bayesian skills.[18]

Testing Analysts

This breakthrough obviously has significant potential meaning for the use of Bayes in intelligence analysis. If natural frequencies allow at least some people to become better Bayesians, maybe such a result could be the foundation for teaching and applying Bayes to entry-level intelligence analysts.

There are substantial differences between doctors and intelligence analysts, however. Doctors are, by and large, steeped in the scientific method and are arguably much more familiar (not to mention comfortable) with mathematics in general and statistics in particular than are most intelligence analysts. In addition, because of the competitive nature and lengthy education involved in becoming a doctor, it is likely that most doctors are older, better educated and more mature than the entry-level analysts that need to be targeted by any Bayesian educational process for intelligence analysis. In short, in order to ensure that natural frequencies might prove useful in educating intelligence analysts about Bayes, it seemed that the first step was to repeat Gigenrenzer's study with these young analysts.

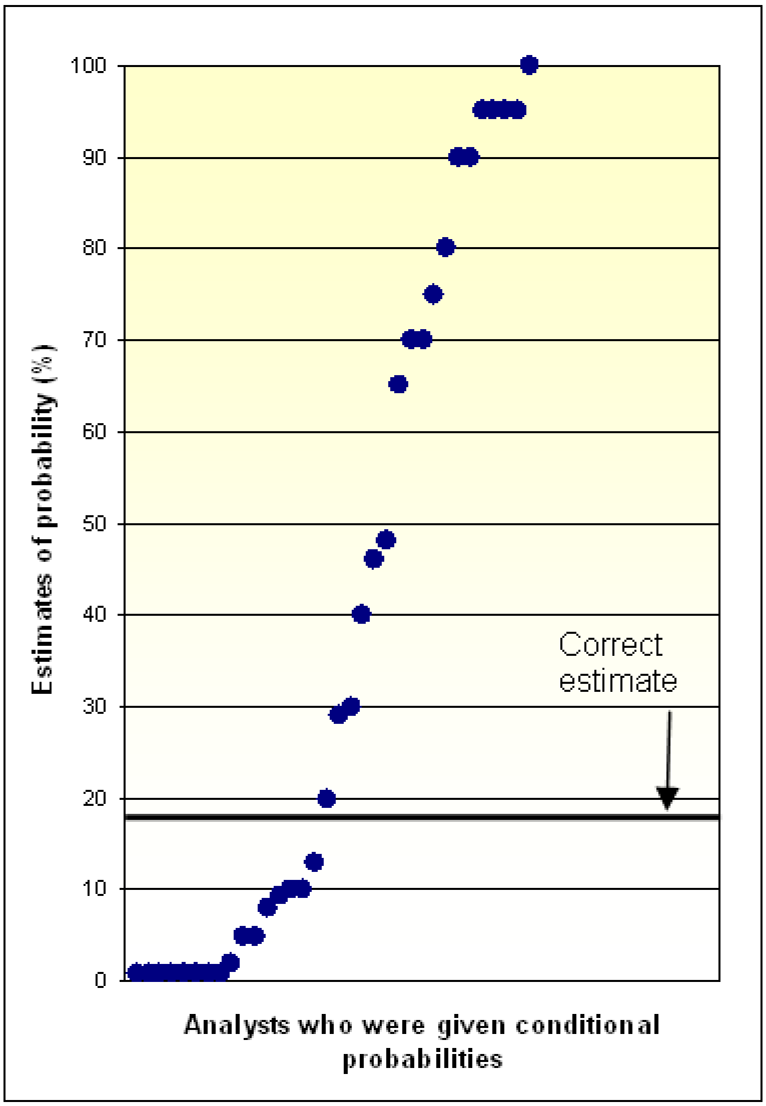
Finding a group of young analysts could have proven difficult. Many analysts require security clearances, and those in classified positions are both difficult to engage as a group and are not typically encouraged to participate in experiments. Fortunately, at Mercyhurst College there exists a large group of potentially testable subjects. Mercyhurst offers the only residential, 4-year degree program in Applied Intelligence Studies in the world. Students participate in a rigorous curriculum (that begins in their first year) designed to prepare them for positions as analysts in the national security, business and law enforcement intelligence communities. Students in the program will normally have participated in numerous real world intelligence projects or simulations of real world intelligence projects and are routinely required to produce high quality open source intelligence products as part of their normal class work. All undergraduates and many graduate students have internships with agencies and organizations within the intelligence community before they graduate. Because of this preparation, virtually all of the graduates of this program are employed within a year by some sector of the intelligence community and, in a typical class, 60–70% of these students will go to work somewhere in the national security community as intelligence analysts. By the time they are seniors, they are, by all relevant measures, experienced intelligence analysts without the clearance.

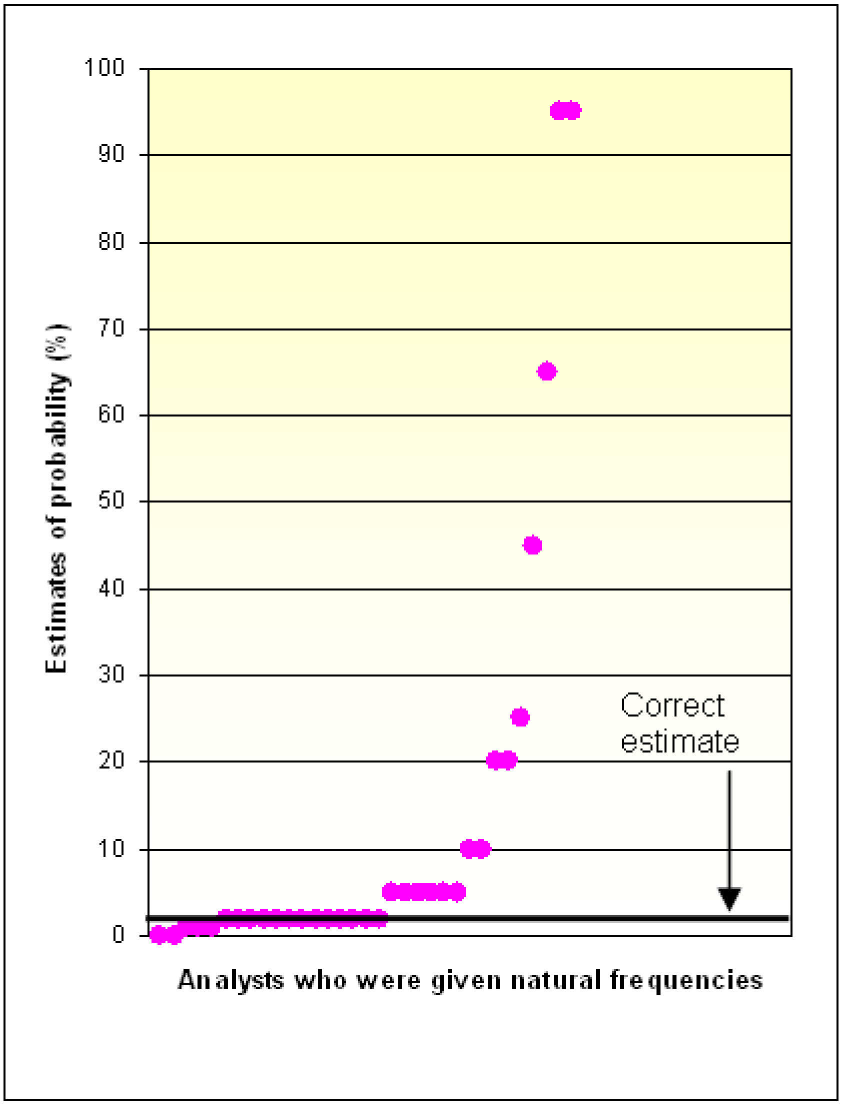
As part of an ordinary classroom exercise, 67 senior intelligence studies students at Mercyhurst were asked to answer two different questions, similar to the ones posed earlier. As with Gigenrenzer's experiment, one used standard statistical terminology (terminology with which all Mercyhurst Intelligence Studies students are familiar as statistics is part of the required course load) while the other group received a similar question using natural frequencies. Unlike Gigenrenzer's experiment, the questions imagined a hypothetical intelligence situation rather than a medical one. The question that utilized standard statistical terminology posed a question about the chances of a country going to war based on a signals intelligence (SIGINT) report, while the natural frequency question asked about the chances of someone being a spy within the intelligence community based on the results of a special test. Both questions were structured identically to Gigenrenzer's questions.

There was a statistically significant (alpha=.05) difference between the traditional statistics group and the natural frequency group. Twenty six of 33 (79%) students fell within 10 points of the correct answer[19] in the natural frequency group while only 6 of 34 (18%) analysts got that close in the

http://scholarcommons.usf.edu/jss/vol2/iss1/3
DOI: http://dx.doi.org/10.5038/1944-0472.2.1.3

traditional statistics group. These results (see charts below) strongly suggest that natural frequencies do, in fact, allow intelligence analysts to be better Bayesians (see statistical annex for additional details). The results from the natural frequencies group tracks exactly (and almost certainly serendipitously) with Gigenrenzer's own findings: 19 of 24 doctors (79%) fell within 10 points of the correct answer.





In order to determine if there were any pre-existing differences between the two subject groups (doctors and analysts) we also compared the doctors who had the problem expressed in traditional statistical terms and the analysts who had a similar problem articulated in traditional statistical terms. While there was a difference in the number of individuals who came within 10 points of the correct answer (33% of doctors vs. 18% of

http://scholarcommons.usf.edu/jss/vol2/iss1/3
DOI: http://dx.doi.org/10.5038/1944-0472.2.1.3

analysts) this difference, contrary to our initial concerns, was not significant (alpha=.05; see statistical annex for additional details), suggesting that analysts who are about to complete the Mercyhurst program are at least as comfortable as doctors with statistically based reasoning.

Conclusion and Next Steps

This report on an ongoing research program suggests that intelligence analysts may well be better able to handle Bayesian style reasoning than expected from previous studies of people in other disciplines if natural frequencies are used. The studies conducted extend, and, to a certain extent, generalize previous findings that natural frequencies are an effective method for encouraging Bayesian reasoning.

If analysts, using natural frequencies, can learn Bayes, as this exercise suggests, what are the next steps? Gigenrenzer himself explored a number of possibilities but they all seem too complicated for day-to-day use by young analysts, some of whom are forward deployed in combat zones. Such analysts need a simple, quick procedure; something that could be done on the back of an envelope and learned in an afternoon. In computerized form, such a program would have to be lightweight and user friendly, perhaps an extension to Mozilla's Firefox[20] or a lightweight piece of specially designed software such as XeroxPARC's ACH 2.0.3.[21]

Fortunately, such efforts are ongoing. One of the co-authors of this article is pursuing thesis research along these precise lines and Mercyhurst has undertaken a joint project with Penn-State Behrend to develop a standalone program that integrates open source information in real-time into a software package that allows analysts to make Bayesian-like decisions on the fly. As Schweizer stated almost 30 years ago, "Bayesian analysis is a method of extracting more information than usual from evidence, and of encouraging analysts to consider alternative explanations of that evidence"[22] and this ongoing research program hopes to bring the benefits of this powerful tool closer to the reach of the entry-level analyst.

Statistical Annex:

1. Is there a significant difference between the Traditional Statistics Group and the Natural Frequencies Group?

Hypothesis Statements:

Null Hypothesis: There is no difference between using the traditional statistics language approach and using the natural frequency approach.

Alternate Hypothesis: There is a significant difference between using the traditional statistics language approach and using the natural frequency approach. (Claim)

Assumption Checking:

For both groups the sample size is large (30 or more).

Independent samples as different approach used for each sample/data set.

Decision is to use Z-distribution two-hypothesis test.

Analysis:

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Frequency per group	List of Answers from Problem using Percentages	34	38.3382	37.78823	6.48063
	List of Answers from Problem using Natural Frequencies	33	13.4606	25.09408	4.36832

http://scholarcommons.usf.edu/jss/vol2/iss1/3
DOI: http://dx.doi.org/10.5038/1944-0472.2.1.3

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Frequency per group	Equal variances assumed	17.876	.000	3.164	65	.002	24.87763	7.86150	9.17713	40.57812
	Equal variances not assumed			**3.183**	57.548	**.002**	24.87763	7.81542	9.23076	40.52450

Z-test value = 3.183

P-value = 0.002

P-value = 0.002 is smaller than level of significance of 0.05, thus reject the null hypothesis.

Conclusion:

At the 5% level, there is a significant difference between using the traditional statistics language approach and using the natural frequency approach.

2. Is there a significant difference between the first Doctor data set and the first Intel Studies data set for the Traditional Statistics Language?

Hypothesis Statements:

Null hypothesis: There is no difference between the Doctor data set and the Intel Studies data set for the Traditional Statistics Language used in Bayesian problems.

Alternative hypothesis: There is a significant difference between the Doctor data set and the Intel Studies data set for the Traditional Statistics Language used in Bayesian problems. (Claim)

Assumption Checking:

Sample size for Doctor data is small (less than 30).

Samples are independent as the data is collected from two different professions.

Normality assumption check:

Null: The variable is normally distributed.

Alternative: The variable is not normally distributed.

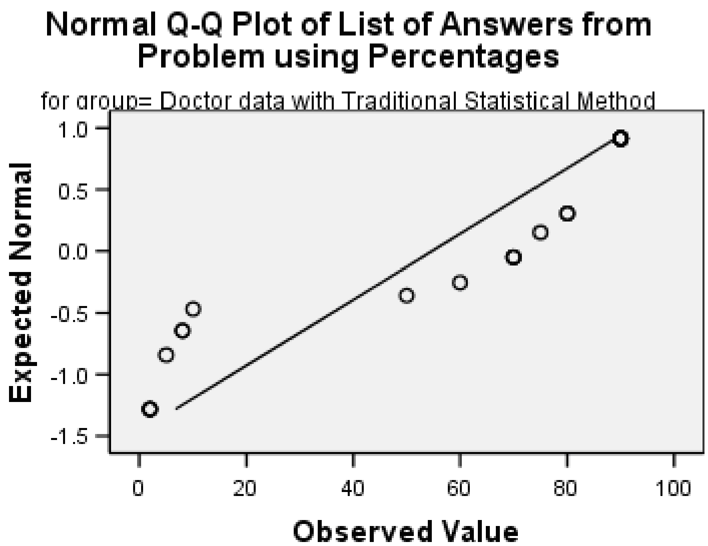
Tests of Normality

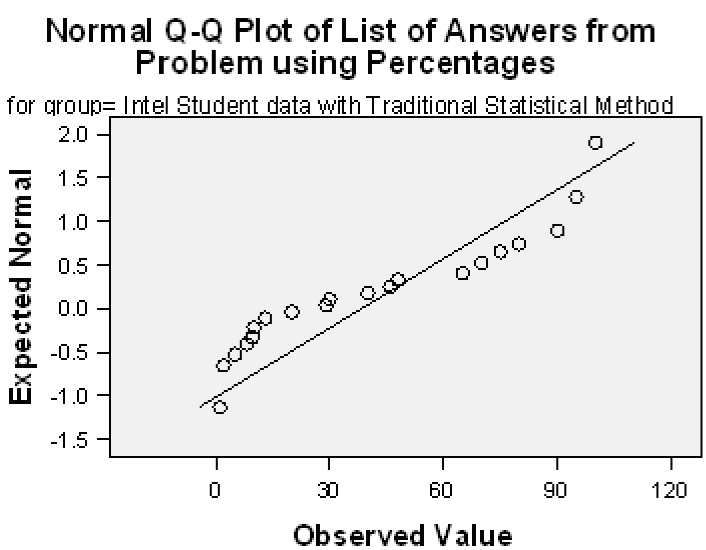
		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Group	Statistic	df	Sig.	Statistic	df	Sig.
List of Answers from Problem using Percentages	Doctor data with Traditional Statistical Method	.241	24	**.001**	.769	24	.000
	Intel Student data with Traditional Statistical Method	.219	34	**.000**	.823	34	.000

a. Lilliefors Significance Correction

From Kolmogorov-Smirnov column from above output table, for both the samples P-values are less than 0.05. Thus reject the null hypothesis. Thus the normality assumption is not satisfied. But normality is robust thus time is to check for the plots.

http://scholarcommons.usf.edu/jss/vol2/iss1/3
DOI: http://dx.doi.org/10.5038/1944-0472.2.1.3


Normal Q-Q Plot of List of Answers from Problem using Percentages
for group= Doctor data with Traditional Statistical Method
Expected Normal
1.0
0.5
0.0
-0.5
-1.0
-1.5
0
20
40
60
80
100
Observed Value


Normal Q-Q Plot of List of Answers from Problem using Percentages
for group= Intel Student data with Traditional Statistical Method
Expected Normal
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
0
30
60
90
120
Observed Value

Both the Normal Q-Q plots show that most of the points are close to the diagonal line. Thus the assumption of normality is satisfied for the samples.

Note that for the doctor data few points (lower left) look as if they are away from the line. But if you look at the vertical axis scale, it really does not matter.

Decision is to use t-distribution independent two-hypothesis test.

Analysis:

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
List of Answers from Problem using Percentages	Doctor data with Traditional Statistical Method	24	54.7500	37.50507	7.65569
	Intel Student data with Traditional Statistical Method	34	38.3382	37.78823	6.48063

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
List of Answers from Problem using Percentages	Equal variances assumed	.010	**.920**	**1.634**	56	**.108**	16.41176	10.04361	-3.70800	36.53153
	Equal variances not assumed			1.636	49.911	.108	16.41176	10.03036	-3.73570	36.55923

http://scholarcommons.usf.edu/jss/vol2/iss1/3
DOI: http://dx.doi.org/10.5038/1944-0472.2.1.3

In order to decide the t-test value, we need to see if the equal variance assumption is satisfied.

Null: The variances are equal.

Alternative: The variances are not equal.

From above output table, according to Levene's test, P-value is 0.920. It is larger than level of significance of 0.05. Thus fail to reject the null hypothesis.

Thus variances are equal at 5% level.

t-test value = 1. 634

P-value = 0.108

P-value = 0.108 is larger than level of significance of 0.05 thus fail to reject the null hypothesis.

Conclusion:

There is no difference between Doctor data set and the Intel Studies data set for the Traditional Statistics Language used in Bayesian problems at 5% level.

About the Authors

Kristan J. Wheaton, Department of Intelligence Studies, Mercyhurst College

Jennifer Lee, Department of Intelligence Studies, Mercyhurst College

Hemangini Deshmukh, Department of Mathematics, Mercyhurst College

Endnotes

1 U.S., *Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction*, Chapter 8, Recommendation 5, http://www.wmd.gov/report/report.html #chapter8 (accessed February 11, 2007).

2 U.S., "Intelligence Reform and Terrorism Prevention Act of 2004," Section 1017 a, http://www.nctc.gov/docs/pl108_458.pdf (accessed February 11, 2007).

3 Michael Schrage, "What Percent Is 'Slam Dunk'?" *Washington Post*, February 20, 2005, http://www.washingtonpost.com/wp-dyn/articles/A37115-2005Feb19.html (accessed February 11, 2007).

4 Richards J. Heuer, ed., *Quantitative Approaches To Political Intelligence: The CIA Experience* (Westview Press, Boulder Colorado, 1978): 4–10.

5 Schrage, op. cit.

6 Bruce Blair, "The Logic of Intelligence Failure - CDI," *Center For Defense Information*, March 9, 2004, http://www.cdi.org/blair/logic.cfm (accessed February 11, 2007).

7 J.J. O'Connor, "Bayes Biography," http://www-groups.dcs.st-andrews.ac.uk/%7Ehistory/Biographies/Bayes.html (accessed February 11, 2007).

8 Peter M. Lee, *Bayesian Statistics: An Introduction* (Arnold Publishers, London, 2004): xiii-xiv. Material in [ ] added by authors.

9 Daniel Kahneman, Paul Slovic and Amos Tversky, *Judgment Under Uncertainty: Heuristics And Biases*, (Cambridge University Press, 1982): 48.

10 Heuer, ed, op. cit., 11–30.

11 Heuer, ed, op. cit., 26.

12 Stephen Losey, "Hiring freeze could deplete FBI staffs," *Federal Times.com*, January 30, 2007, http://federaltimes.com/index.php?S=2515739 (accessed February 11, 2007).

13 Peter Sedlmeier and Ger Gigenrenzer, "Teaching Bayesian Reasoning In Less Than Two Hours," *Journal Of Experimental Psychology*, Vol 130, No. 3 (2001): 382–384.

14 Heuer, ed., op. cit., 14–15.

15 Peter Sedlmeier and Gerd Gigenrenzer, "Teaching Bayesian Reasoning In Less Than Two Hours," *Journal Of Experimental Psychology*, Vol. 130, No. 3 (2001): 382.

16 Robert J. Sternberg, *Cognitive Psychology*, Thomson Wadsworth, Belmont, California (2006): 451–452.

17 Sedlmeier and Gigenrenzer, op. cit.

18 Ulrich Hoffrage and Gerd Gigerenzer, "Using Natural Frequencies to Improve Diagnostic Inferences," *Academic Medicine* 73 (1998): 538–540.

19 The intent of the experiment was to test reasoning skills and not the ability to do mathematics. A ten-point margin seemed that it would account for any math errors while still indicating a reasoning style.

20 Mozilla's Firefox is a powerful, free web browser that allows user to add multiple "extensions" that alter and customize the web browsing experience. The software is available at http://www.mozilla.com/en-US/firefox/.

http://scholarcommons.usf.edu/jss/vol2/iss1/3
DOI: http://dx.doi.org/10.5038/1944-0472.2.1.3

21 ACH 2.0.3 is a lightweight, user-friendly piece of software developed by XeroxPARC and designed to partially automate the Analysis Of Competing Hypotheses method of intelligence analysis. The software is available for download from http://www2.parc.com/istl/projects/ach/ach.html.

22 Heuer, op. cit., 11.

http://scholarcommons.usf.edu/jss/vol2/iss1/3
DOI: http://dx.doi.org/10.5038/1944-0472.2.1.3