8-2002

Coverage Properties of the Inverse Sinh Transformation and the Adjusted Wald Confidence Intervals for the Odds Ratio and the Relative Risk.

Troy Allen Bowman
East Tennessee State University

Coverage Properties of the Inverse Sinh Transformation and the Adjusted Wald

Confidence Intervals for the Odds Ratio and the Relative Risk

---

A Thesis

Presented to the Faculty of the Department of Mathematics

East Tennessee State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Mathematical Sciences

---

by

Troy Allen Bowman

August 2002

Robert Price, Jr., Chair

Edith Seier

T. Henry Jablonski, Jr.

Keywords: Confidence Interval, Relative Risk, Odds Ratio

ABSTRACT

Coverage Properties of the Inverse Sinh Transformation and the Adjusted Wald

Confidence Intervals for the Odds Ratio and the Relative Risk

by

Troy Allen Bowman

The inverse sinh transformation on the Woolf interval is used to calculate the confidence interval for the odds ratio and the relative risk in a 2×2 table. According to Robert Newcombe, the new interval should improve the coverage probabilities and shorten the width of the confidence interval for these ratios, but the new interval requires evaluation of coverage properties. In this thesis, we will evaluate the exact coverage properties of this modified interval in extreme cases. Also , we will compare the coverage properties of this new interval to other widely-used adjusted intervals. Through comparisons of exact coverage probabilities and interval widths, we will discover if Newcombe's inverse sinh transformation provides better coverage properties than the adjusted methods.

DEDICATION

This thesis is dedicated to my family, who has supported my efforts to complete my graduate degree. Thanks for your love and support.

ACKNOWLEDGEMENTS

A special thanks to my thesis advisor, Dr. Robert Price Jr., who has been so helpful and patient with me throughout this process. Thanks so much for your support.

Contents

List of Tables

CHAPTER 1

Introduction

Computing confidence intervals for ratios such as the odds ratio and relative risk are important inferential procedures in the statistical analysis of categorical data. However, it is easy to find and implement numerous large sample confidence intervals for a particular set of data. Therefore, the coverage probabilities for approximate confidence intervals are of strong interest in both small and large samples. If a proposed confidence interval of $1 - \alpha$ does not achieve this intended coverage probability, then problems arise in the relevance of the conclusions drawn from the data based on this confidence interval. Also, if the confidence interval is very wide, then the conclusions drawn from the confidence interval are not relevant either. Note, a confidence interval has a lower bound and an upper bound. The width or length of a confidence interval refers to the difference between the upper bound and the lower bound. When dealing with confidence intervals, the researcher always wants high confidence and short intervals. Hence, the researcher must decide on the formula to use for calculating the confidence interval for different data sets. Also, one must remember that each formula for calculating the confidence interval has different coverage probabilities and different interval widths. If the sample size is large, then the intervals provide approximately the same high coverage probabilities along with short interval widths. What happens when sample size is very small, the cell frequency of a single cell is very small, or when the probabilities of certain outcomes within the data set are very small as well? In these extreme cases, the confidence interval can produce overshoot, or the confidence interval may not even be able to be calculated at all. Also, some intervals may be conservative or even liberal in certain situations. By being liberal, the confidence interval does not achieve the level of confidence $1-\alpha$ exactly, but the coverage probability is less than $1-\alpha$ level stated for the test. While being conservative implies that the confidence level of $1-\alpha$ is not only reached, but that coverage property is exceeded. An even higher level confidence interval is formed. What about the length of the newly formed confidence interval? The ideal confidence interval will not only provide the $1-\alpha$ coverage probability desired, but the confidence interval will also have the shortest width. It is easy to obtain a high coverage probability for a confidence interval if the width of the interval is not considered. In theory, the confidence interval $[0, \infty]$ will have the desired coverage probability, but with the large width, this confidence interval is otherwise useless. A large width can be illustrated with liberal confidence intervals. These intervals have high coverage probabilities, but they are usually very wide as well. A very undesired property for a confidence interval. Therefore, when new formulas for calculating the confidence interval are proposed, a study of the coverage probabilities and the width of these new intervals must be undertaken, and the results clearly defined.

One proposed confidence interval for the odds ratio is the inverse sinh transformation on the Woolf interval [1]. When measuring the odds ratio in an ordinary 2×2 table, the odds ratio is $\omega = \frac{ad}{bc}$ where a, b, c, and d are the four cell frequencies in the 2×2 table. Consider the Woolf interval $\ln \omega \pm y$, where $y = z\sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$. This interval has good coverage probabilities and short interval widths when sample size is large and cell entries are equivalent, but in extreme cases when n is small or cell entries zero, the coverage probabilities falter. Hence, apply the inverse sinh transformation to form the new confidence interval of the form $\ln \omega \pm 2x$, where $x = \sinh^{-1}(\frac{y}{2})$[1], and discover if this transformation improves the coverage probabilities and shortens the widths of the confidence intervals formed in extreme cases.

Now, consider another ratio in statistics, the relative risk. For an ordinary 2×2 table with cell frequencies a,b,c, and d, the relative risk is $RR = \frac{(\frac{a}{n_1})}{(\frac{b}{n_2})}$ with $n_1 = a + c$ and $n_2 = b + d$. Also, n_1 and n_2 are fixed and can be thought of as two independent groups. The Woolf confidence interval for relative risk is $\ln RR \pm y$ where $y = z\sqrt{\frac{1}{a} + \frac{1}{b} - \frac{1}{n_1} - \frac{1}{n_2}}$. This interval, as seen before with the odds ratio, also falters in the extreme cases of a small sample size or cell entries a or $b = 0$. Therefore, consider the inverse sinh transformation on the Woolf interval as $\ln RR \pm \sinh^{-1}(\frac{y}{2})$[1] in order to study the coverage probabilities of this new interval for any improvements in coverage probabilities and interval widths of the confidence intervals formed in extreme cases.

CHAPTER 2

Odds Ratio

Measures of association are used to measure the strength of an association between variables. One such measure of association is the odds ratio. The odds deal with the probability of a yes outcome in a population. An odds ratio compares the odds of the yes probability in one group with the odds of the yes probability in the other group. The odds ratio can range from values of 0 to ∞. An odds ratio equal to 1 implies that there is no association between the two groups, or the odds are equal between the two groups. For an odds ratio greater than 1, Group 1 has a larger chance to have a yes outcome than Group 2. Group 1 has greater odds of occuring than Group 2. Now for an odds ratio less than 1, Group 1 has less chance to occur than Group 2. Group 2 has the greater odds of occuring than Group 1. The odds ratio is widely used in medical research, for example, to compare treatment effects of the two groups in order to find if the odds of a success are the same for both groups. But what about the difference in proportions? Differences between two sample proportions can also be used to compare treatment effects of two groups. However, the differences tend to vary in meaning when proportions are near 0 or 1. If p_1 is the probability of a disease in the absence of treatment, and p_2 is the probability of a disease under treatment with a preventive drug, then a difference of .05 may be rather small if the true probabilities are .5 and .45. The same difference, however, can have considerable practical importance if the true probabilities are .10 and .05[2]. In the second case, the disease could have been prevented in 1 out of 2 people who contracted the disease, a very significant result from a small difference. This illustrates the use of the difference between two proportions may not be the best way to compare the two groups, because a small difference can be very misleading in respect to the overall situation. Since the odds ratio seems to be the better way to compare two proportions, confidence intervals for the odds ratio are of particular importance along with the coverage probabilities and the width of these confidence intervals. The odds ratio for an ordinary 2×2 table is $\omega = \frac{ad}{bc}$ with a,b,c, and d as the four cell frequencies. If all the cell frequencies are approximately the same or all quite large, then any formula used to calculate the confidence interval for the odds ratio works very well. That is, all confidence intervals have good coverage properties and similar, short widths in this situation. What happens when the population n is small, or some cell frequencies are very small or even zero? For these small sample sizes, the sampling distribution is highly skewed. The use of the natural logarithm is implemented to the odds ratio, since the natural log has less skewness. This is illustrated by the formulas chosen to calculate the confidence intervals for the odds ratio. However, these extreme cases can still have a huge effect on the coverage probabilities and the widths of the confidence intervals for the odds ratio.

First, let's consider an example to illustrate the calculation of an odds ratio and the confidence interval for the odds ratio. Some say that vitamin C can help prevent the common cold. Hence, a Canadian experiment examined the claim with the results in Table 1 [4]. From the table, an estimate of the odds ratio is $\omega = \frac{(335)(105)}{(76)(302)} = 1.533$. This means that the odds of catching a cold taking the placebo are 1.53 times higher than the odds of catching a cold taking vitamin C. The same as saying the odds

Table 1: Vitamin C and the common cold

	Cold	No Cold	Totals
Placebo	335	76	411
Vitamin C	302	105	407
Totals	637	181	818

of catching a cold taking the placebo are 53.3% greater than the odds of catching a cold taking vitamin C. An approximate 95% confidence interval for the log odds ratio is $\ln(1.53) \pm 2(\sinh^{-1}(\frac{1.96}{2}(\sqrt{\frac{1}{335}+\frac{1}{76}+\frac{1}{302}+\frac{1}{105}}) = .4269 \pm .3321$ or $(.0948, .759)$. Therefore, an approximate 95% confidence interval for the odds ratio is $(e^{.0948}, e^{.759})$ which equals $(1.099, 2.136)$. The research indicates with 95% confidence that the true odds of catching a cold taking the placebo are 1.099 *to* 2.136 times higher than catching a cold taking the vitamin C.

Now, what happens when calculating the odds ratio in extreme cases? An exact confidence interval for the odds ratio can be calculated. This exact interval, however, is very hard to calculate and gives a resulting confidence interval that is liberal and very wide. Therefore, consider the inverse sinh transformation on the Woolf's interval as $\ln\omega \pm 2x$, where $x = \sinh^{-1}(\frac{y}{2})$ and $y = z\sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}$ [1]. The confidence interval works well in normal cases, but what happens with the coverage probabilities and the width of the interval in extreme cases? How will the coverage probabilities and the width of the inverse sinh transformation confidence interval be effected in extreme cases as compared to other confidence intervals for the odds ratio? One such widely-used confidence interval is an adjustment to the Woolf interval by the addition of $\frac{1}{2}$ to each cell frequency. The adjustment interval is the $\ln\omega \pm y$ where $\omega =$

$\frac{(a+.5)(d+.5)}{(b+.5)(c+.5)}$ and $y = z\sqrt{\frac{1}{a+.5}+\frac{1}{b+.5}+\frac{1}{c+.5}+\frac{1}{d+.5}}$. Will the adjustment interval work as well or better in extreme cases as compared to the inverse sinh transformation?

To do the comparisons and to define the coverage properties for the inverse sinh transformation, a MatLab program (Appendix B : Program 1) is used. With the program, changing the population size n and randomly selecting cell frequencies for n can be computed easily. The first step in the program is to generate a matrix of all possible cell frequencies for a given population n. The process is referred to as multinomial sampling, since the total population is fixed but not the distribution in the cell frequencies. Now the inverse sinh transformation, unlike the Woolf interval, can deal with a single cell entry of zero. Therefore, the matrix of all possible cell frequencies must be altered in such a way that only the cell frequencies with at most one zero cell entry remain. For example, if cell entry $a = 0$ and by letting $a \to 0$, holding b, c, and d fixed, then

$$\ln U = \ln \omega + 2x$$

$$\sim \ln U = \ln \tfrac{d}{bc} - \ln(\tfrac{4sinh^2 x}{z^2} - \tfrac{1}{b} - \tfrac{1}{c} - \tfrac{1}{d}) + 2x$$

$$\sim \ln \tfrac{z^2 d}{4bc} + 2x - 2\ln\sinh x$$

$$U \sim \tfrac{z^2 d}{4bc}(\tfrac{e^x}{\sinh x})^2$$

$$\to \tfrac{z^2 d}{bc}$$

[1].

Also, for the upper limit when $d = 0$ and the lower limit when $b = 0$ and $c = 0$, substitute z^2 into that cell entry[1]. When $d = 0$ the interval $[0, \frac{az^2}{bc}]$ serves as the confidence interval for the odds ratio. Also when $b = 0$, the interval $[\frac{ad}{z^2 c}, \infty]$ serves as the confidence interval.

After the calculation of the confidence intervals, the program focuses on the coverage probabilities for the given confidence interval, labeled *covadj* and *covnew* with *covnew* standing for the inverse sinh transformation interval. The program can also randomly select probabilities for each cell entry to calculate the odds ratio, or values can be entered to produce a desired odds ratio, for instance 1, 2, 4, etc. Next, the program calculates the coverage probabilities of the given intervals under examination. By changing values of n or the probability of each cell entry, the effect on the coverage probabilities of the inverse sinh transformation of the Woolf's interval can be examined. This examination can provide information on the reaction of the confidence interval and comparisons between the inverse sinh transformation and the adjusted confidence interval for the odds ratio. Also notice, finding the interval widths for the two methods can be accomplished easily in MatLab. This will lead to another comparison between the two confidence intervals. Therefore, looking at the coverage probabilities and the widths of the two confidence intervals will provide information on which confidence interval has the best coverage properties in extreme cases.

CHAPTER 3

Relative Risk

The odds ratio is a useful measure of association regardless of the method used to collect the data. However, the odds ratio has special meaning for cross-sectional and prospective studies, because it is used to estimate a quantity called relative risk [3]. The key to calculate relative risk for the two types of studies is a fixed sample size n_1 and n_2 for the explanatory variable. Only when n_1 and n_2 are fixed, can one calculate the relative risk. The relative risk is referred to in principle as the ratio of the success probabilities for two groups, $\frac{p_1}{p_2}$. A difference between these two probabilities may have great importance when one or both probabilities are near 0 or 1. Like the odds ratio, relative risk can range from values of 0 to ∞. A relative risk equal to 1, occurring when $p_1 = p_2$, means the probability of success is the same for both groups and the response is independent to a group. Relative risk acts the same as the odds ratio when it is greater than 1 and less than 1. In practice, the relative risk is a parameter of major importance in the medical field, because it is referred to as the risk of developing a condition, usually a disease, for one group compared to another group. Researchers have many different methods to calculate confidence intervals for the relative risk, but researchers want to use the most efficient method to make calculations. That is, the method that offers the best coverage probabilities and shortest interval width for the confidence interval of relative risk. When dealing with a large population of the two groups, n_1 and n_2, and with p_1 and p_2 approximately equal, .5, it follows that any formula used for calculating the confidence interval for relative risk would have great coverage probabilities and short interval widths. However, when n is small and one of the probabilities is near 0 or 1, the coverage probabilities and interval widths for a given confidence interval can be effected greatly. With a small sample size the relative risk can be highly skewed as well, so the natural log transformation is used when calculating confidence intervals for relative risk.

Let's look at an example in order to calculate the relative risk. There seems to be an association between aspirin and heart attacks. A research group did a five year study testing whether regular intake of aspirin reduces mortality from cardiovascular disease (MI) with the following results [3]. From the data, no-

Table 2: Aspirin Use and Myocardial Infarction (MI)

Group	MI Yes	MI No	Total
Placebo	189	10845	11034
Aspirin	104	10933	11037

tice two independent binomial samples of size $n_1 = 11,034$ and $n_2 = 11,037$. An estimate of the relative risk is $\frac{(\frac{189}{11034})}{(\frac{104}{11037})} = 1.818$. This means that the proportion of heart attacks (MI cases) was 81.8% higher for the group of doctors taking the placebo. An approximate 95% confidence interval for the log relative risk is $\ln(1.818) \pm 2(\sinh^{-1}(\frac{1.96}{2}(\sqrt{\frac{1}{189} + \frac{1}{104} - \frac{1}{11{,}034} - \frac{1}{11{,}037}}) = .5976 \pm .2373$ or $(.3603, .8349)$. The approximate 95% confidence interval for the relative risk is $(e^{.3603}, e^{.8349})$ which equals $(1.434, 2.305)$. The research indicates with 95% confidence that the true proportion of heart attack cases for the group taking the placebo is between 1.434 to 2.305 times the proportion of heart attacks for the group taking aspirin. Illustrating the risk of a heart attack is at least 43.4% higher for the group taking the placebo. This is a very important result. What would happen if the researchers looked at the difference in proportions instead of relative risk? The approximate 95% confidence interval for the difference in proportions is found to be $(.005, .011)$. This difference is very small and it seems like the two groups are not very different at all. The relative risk, however, illustrates a major difference that is important to public health. So, when the proportions are very small the difference of the proportions can be very misleading as illustrated in this example.

Since relative risk and the odds ratio are very similar, calculating an exact confidence interval for the relative risk can also be considered. However, this process is not only long and hard to do in general, but the results are a liberal confidence interval that is very wide in length. Consider the inverse sinh transformation on the Woolf's interval, since it is easy to calculate and may give shorter interval widths. Remember, this interval is $\ln RR \pm 2\sinh^{-1}(\frac{y}{2})$ where $y = z\sqrt{\frac{1}{a} + \frac{1}{b} - \frac{1}{n_1} - \frac{1}{n_2}}$[1] as noted before. Will this transformation perform well in extreme cases? Other formulas used to calculate the confidence interval for relative risk deal with adjustments on the Woolf's interval. Adjustment 1 deals with the addition of $\frac{1}{2}$ of a success to the sample implying the confidence interval $\ln RR \pm y$ with $RR = \frac{(\frac{(a+.5)}{n_1})}{(\frac{(b+.5)}{n_2})}$ and $y = z\sqrt{\frac{1}{a+.5} + \frac{1}{b+.5} - \frac{1}{n_1} - \frac{1}{n_2}}$. Adjustment 2 deals with the addition of $\frac{1}{2}$ of a success and $\frac{1}{2}$ of a failure to the sample implying that the confidence interval is $\ln RR \pm y$ with $RR = \frac{(\frac{(a+.5)}{(n_1+1)})}{(\frac{(b+.5)}{(n_2+1)})}$ and $y = z\sqrt{\frac{1}{a+.5} + \frac{1}{b+.5} - \frac{1}{n_1+1} - \frac{1}{n_2+1}}$. Therefore, comparing the coverage probabilities and the widths of the confidence intervals will reveal if the new inverse sinh transformation provides better coverage probabilities and shorter interval widths than the widely-used adjusted intervals.

In order to examine the coverage probabilities, interval widths, and to make comparisons between the confidence intervals, the use of a computer program is needed. The MatLab program (Appendix B : Program 2) can be used to sample different population sizes n_1 and n_2, along with different probabilities p_1 and p_2, in order to study the coverage properties of these confidence intervals. By changing sample size, cell frequencies, and probabilities, the effects of the extreme cases on coverage probabilities and the width of these confidence intervals can be examined.

The relative risk for an ordinary 2×2 table having cell frequencies a,b,c, and d with $n_1 = a + c$ and $n_2 = b + d$ is $RR = \frac{(\frac{a}{n_1})}{(\frac{b}{n_2})}$. The program utilizes independent binomial sampling, since n_1 and n_2 are fixed. Meaning when there are two categories, we assume a binomial distribution for the sample in each column, with the number of trials equal to some fixed column total. Next, the program builds a matrix corresponding to all possible combinations of values for cell frequencies a and b. The inverse sinh transformation of the Woolf confidence interval for relative risk deals with a single cell entry of zero as well. Therefore, the removal of any possible combination of cell frequencies when both a and b are zero must be completed. The adjusted formulas, however, can deal with multiple cell frequencies of zero. For the inverse sinh transformation, the method copes with a or b equal to zero by substituting z^2 for the zero cell entry [1]. So as $a \to 0$, holding b fixed, let

$$\ln U = \ln RR + y$$

$$\sim \ln U = \ln \frac{\frac{1}{n_1}}{\frac{b}{n_2}} - \ln\left(\frac{4sinh^2 x}{z^2} - \frac{1}{b} - \frac{1}{n_1} - \frac{1}{n_2}\right) + y$$

$$\sim \ln \frac{\frac{z^2}{n_1}}{\frac{4b}{n_2}} + y - 2 \ln \sinh y$$

$$U \sim \frac{\frac{z^2}{n_1}}{\frac{4b}{n_2}}\left(\frac{e^x}{\sinh x}\right)^2$$
$$\to \frac{\frac{z^2}{n_1}}{\frac{b}{n_2}}$$

.

As seen with the odds ratio, when $a = 0$ the interval $[0, \frac{\frac{z^2}{n_1}}{\frac{b}{n_2}}]$ serves as the confidence interval for relative risk. Also when $b = 0$, the interval $[\frac{\frac{a}{n_1}}{\frac{z^2}{n_2}}, \inf]$ serves as the confidence interval for relative risk.

Next, after calculating the confidence intervals for relative risk, find the coverage probabilities for these intervals, labeled *covadjwoolf* and *covnew* in the program. *Covnew* stands for the coverage probabilities for the inverse sinh transformation interval, and *covadjwoolf* stands for the coverage probabilities for the adjusted Woolf intervals. The program also easily calculates the interval widths for the confidence intervals along with random assignment or deliberate assignment of probabilities p_1 and p_2 used to calculate the coverage probabilities for the confidence intervals. Now changing the values of n_1, n_2, p_1, and p_2 illustrates the effect of the extreme cases on coverage probabilities and interval widths. Studying these effects of the extreme cases on the coverage probabilities and the widths of the confidence interval will lead to the effectiveness of the inverse sinh transformation confidence interval as compared to the other well-known adjusted confidence intervals.

CHAPTER 4

Results

The suggested modified inverse sinh transformation confidence interval for the odds ratio and the relative risk requires evaluation of coverage properties to well-known adjusted confidence intervals of the odds ratio and the relative risk. In order to compare the coverage properties of these confidence intervals, two properties of these intervals were examined in this thesis. The properties considered are the coverage probabilities and the width of the confidence intervals at different sample sizes, at all possible cell frequencies in the 2×2 table with these sample sizes, and at a given $1-\alpha$ level of confidence. Sample sizes were extreme in order to analyze the coverage properties of the intervals in extreme cases, since most intervals have similar coverage properties when sample size is large.

The first property used to compare the confidence intervals was the coverage probabilities for a given confidence interval at a $1-\alpha$ level of confidence. The coverage probabilities include the minimum coverage, mean coverage, and the mean squared error of coverage probabilities for all possible cell frequencies of the 2×2 table with sample size n. Remember if we are calculating a level C confidence interval, the confidence interval formed should have coverage probabilities very close to this level C confidence. If not, the interval formed is not very useful. This thesis examines the mean coverage and the minimum coverage of the confidence interval at different sample sizes and $1-\alpha$ levels of confidence. Hopefully, the mean coverage for the confidence intervals will be the same as the $1-\alpha$ level of confidence, and the minimum coverage will be very close to the $1-\alpha$ level of confidence as well. The programs in the thesis also calculate the mean squared error for the probability coverage about the nominal level. The mean squared error (MSE) provides an estimate for the variation in the coverage probabilities for the confidence intervals formed at this $1-\alpha$ level of confidence. Since the mean square error measures the variability in the coverage probabilities, hopefully this value is small in order to illustrate a small variation in the coverage probabilities for the confidence intervals formed at a given $1-\alpha$ level of confidence and at a given sample size n.

The other coverage property examined in this thesis is the width of the confidence interval at a given sample size n. Recall, the width of a confidence interval is equal to the difference between the upper bound and the lower bound of the formed confidence interval. In the program, the interval widths were calculated for every possible cell frequency of a given sample size n. By looking at the five number summary of the widths of the confidence intervals, the summary will be useful in comparing the width of these intervals at a given sample size n. Note, a good confidence interval will have high confidence and short interval width. How will the coverage properties for the inverse sinh transformation confidence interval compare to the coverage properties of the widely-used adjusted confidence interval for the odds ratio and relative risk?

Through the use of the MatLab programs, the coverage properties for the inverse sinh transformation confidence interval and the adjusted confidence intervals were calculated and recorded in the tables found in Appendix A of the thesis. To compare the coverage probabilities of the confidence interval for the odds ratio using the two methods, different confidence intervals were formed at different levels of confidence

(90%, 95%, and 99%). At each of these levels of confidence, different sample sizes were used along with a maximum odds value in order to compare the coverage probabilities of the inverse sinh transformation confidence intervals for odds ratio with the adjustment method confidence intervals for odds ratio. The maximum odds value refers to a limit placed on the odds ratio ω in order to keep the odds ratio between ω and $\frac{1}{\omega}$ before calculating the coverage probabilities. This was implemented in order to see if one of the methods for calculating the confidence intervals is effected by large values of the odds ratio. Looking at the tables of coverage probabilities for the confidence interval for the odds ratio, Tables 3-5, one notices a trend. At any level of confidence and at any sample size, the mean coverage of the confidence interval for the adjustment method is always better, closer to the level of confidence, than the inverse sinh method (Newcombe). Also notice that the minimum coverage and the MSE almost always follow the same trend. In only a few cases, with a maximum odds value ≥ 16, the inverse sinh transformation has a larger minimum coverage probability and a smaller MSE than the adjustment method. Therefore from the tables illustrating the coverage probabilities, it seems that the adjustment method offers better coverage probabilities for the confidence interval of the odds ratio than the inverse sinh transformation method. The inverse sinh transformation only performs as well as the adjustment method is cases where the odds ratio takes on a large value. What about the interval width? Remember that a confidence interval should have a high coverage probability but a short width. By calculating the five number summary for the width of the confidence intervals formed with the two methods, Table 9, notice that the trend continues. At different sample sizes, the confidence intervals formed by the adjustment method have a shorter widths than do the confidence intervals formed by the inverse sinh transformation. This can be easily illustrated from Table 9. Remember, the program is finding the width of all possible confidence intervals formed from all possible cell frequencies at a given sample size n. Notice the differences in the sizes of the interquartile ranges for the two five number summaries. The adjustment method has a smaller interquartile range for all sample sizes n. We can conclude that 50% of the interval widths fall in this range, so the adjustment method does perform better. The table reveals a smaller width for the confidence intervals formed by the adjustment method as compared to the inverse sinh transformation at a given sample size n. Hence, with higher coverage probabilities and shorter interval widths, the adjustment method performs better and will provide a better estimate for a confidence interval for the odds ratio than the inverse sinh transformation.

What about the relative risk? As with the odds ratio, a MatLab program was used, and the results of the coverage probabilities of the confidence intervals for relative risk were recorded in Tables 6-8. As before, the same levels of confidence (90%,95%, and 99%) were used along with different sample sizes, n_1 and n_2. Also, a maximum relative risk of 99 was used in order to calculate the coverage probabilities for the confidence intervals using the three methods. These three methods are the inverse sinh transformation (Newcombe), the method of adding $\frac{1}{2}$ of a success (Adjustment 1), and the method of adding $\frac{1}{2}$ of a success and $\frac{1}{2}$ of a failure (Adjustment 2). From the tables, notice Adjustment 2 exhibits great coverage probabilities even in very extreme cases. The mean coverage for Adjustment 2 is almost the same as the given confidence level, and the minimum coverage is very high as well. Also, the MSE for Adjustment 2 is very small, which means a small variability in the coverage probabilities, another plus. Therefore, Adjustment 2 gives the best coverage probabilities, while the inverse sinh transformation seems to have the worst coverage probabilities of the three methods. Now, the attention turns to the width of the confidence intervals formed by the three methods. Notice from Table 10, the widths of the confidence intervals were calculated at the different sample sizes. The interquartile ranges for the widths of the three methods are almost identical. It seems that all three methods result in a confidence interval with very similar widths. With the maximum values only occuring at a small probability, the inverse sinh transformation can compete with the other two methods when looking at the interval width alone. However, with the higher coverage probabilities and similar interval widths, Adjustment 2 would be the best method to use in order to calculate the confidence interval for the relative risk in extreme cases.

What do the results tell us about the inverse sinh transformation? We have decided that Newcombe's method does not perform as well as other adjustment methods in extreme cases. The first limitation of the inverse sinh transformation is dealing with only one cell frequency equal to zero. The adjustment method on the other hand can deal with multiple cell frequencies of zero. Also, the adjustment methods offer better coverage probabilities and shorter interval widths for the confidence intervals of relative risk and the odds ratio as compared to the inverse sinh transformation. This does not mean, however, that the inverse sinh transformation is useless, only that it has limitations. The key is to know these limitations in extreme cases. This knowledge can aid researchers in the decision to choose an appropriate method for calculating the odds ratio and relative risk to interpret data more effectively in a $2x2$ table.

BIBLIOGRAPHY

[1] Robert G. Newcombe, Logit Confidence Intervals and the Inverse Sinh Transformation, *The American Statistician* **55** (2002), 200-202.

[2] F.L. Ramsey and D.W. Schafer, *The Statistical Sleuth*, Duxbury Press, Belmont, CA, 1997.

[3] Alan Agresti, *An Introduction to Categorical Data Analysis*, John Wiley and Sons, New York, 1996.

[4] T.W. Anderson and D.B. Reid,Vitamin C and the Common Cold, *Canadian Medical Association Journal* **70** (1972), 503-508.

[5] David S. Moore, *The Basic Principles of Statistics*, W.H. Freeman and Company, New York, NY, 1995.

APPENDIX A: TABLES

Table 3: Exact Coverage Probabilities of 90% Confidence Intervals for Odds Ratio

		Newcombe			Adjustment		
n	Max Odds	Min Cov	Mean Cov	MSE	Min Cov	Mean Cov	MSE
10	2.25	.0067	.5773	.1652	.0074	.6580	.1406
20	2.25	.0026	.7249	.0749	.0027	.7849	.0712
30	2.25	.0141	.7836	.0492	.0167	.8343	.0447
10	5.44	.0003	.5770	.1672	.0003	.6505	.1446
20	5.44	.0060	.7259	.0817	.0067	.7858	.0729
30	5.44	.0002	.7827	.0511	.0002	.8311	.0465
10	16	.0003	.5780	.1689	.0000	.6383	.1496
20	16	.0015	.7043	.0988	.0019	.7550	.0900
30	16	.0052	.7740	.0606	.0052	.8172	.0564
10	81	.0002	.5498	.1934	.0002	.5913	.1794
20	81	.0000	.5409	.2027	.0001	.5830	.1888
30	81	.0000	.7643	.0664	.0003	.7991	.0683
10	∞	.0000	.5063	.2318	0	.5297	.2355
20	∞	.0000	.6786	.1187	0	.6962	.1274
30	∞	.0000	.7372	.0912	0	.7467	.0996

Table 4: Exact Coverage Probabilities of 95% Confidence Intervals for Odds Ratio

		Newcombe			Adjustment		
n	Max Odds	Min Cov	Mean Cov	MSE	Min Cov	Mean Cov	MSE
10	2.25	.0015	.6226	.1746	.0015	.6824	.1541
20	2.25	.0024	.7847	.0767	.0024	.8312	.0695
30	2.25	.0177	.8398	.0494	.0184	.8775	.0452
10	5.44	.0001	.6090	.1892	.0001	.6607	.1725
20	5.44	.0183	.7747	.0849	.0197	.8183	.0779
30	5.44	.0144	.8294	.0587	.0154	.8650	.0546
10	16	.0009	.6055	.1939	.0009	.6496	.1786
20	16	.0011	.7492	.1019	.0011	.7822	.0944
30	16	.0001	.8230	.0625	.0001	.8568	.0582
10	81	.0000	.5731	.2227	.0000	.6080	.2084
20	81	.0001	.7479	.1035	.0002	.7783	.0976
30	81	.0000	.8117	.0698	.0000	.8386	.0664
10	∞	.0000	.5434	.2502	0	.5598	.2532
20	∞	.0000	.7049	.1424	0	.7178	.1488
30	∞	.0000	.7716	.1014	0	.7832	.1084

Table 5: Exact Coverage Probabilities of 99% Confidence Intervals for Odds Ratio

		Newcombe			Adjustment		
n	Max Odds	Min Cov	Mean Cov	MSE	Min Cov	Mean Cov	MSE
10	2.25	.0050	.6739	.1744	.0052	.7003	.1657
20	2.25	.0107	.6789	.1699	.0109	.7052	.1579
30	2.25	.0162	.8722	.0582	.0175	.8880	.0559
10	5.44	.0001	.6568	.1892	.0001	.6787	.1817
20	5.44	.0131	.8310	.0866	.0130	.8115	.0906
30	5.44	.0124	.8791	.0522	.0131	.8936	.0498
10	16	.0018	.6369	.2063	.0018	.6577	.1982
20	16	.0009	.7865	.1125	.0009	.8043	.1088
30	16	.0022	.8582	.0649	.0022	.8729	.0626
10	81	.0001	.6033	.2379	.0001	.6234	.2287
20	81	.0014	.7869	.1123	.0014	.8037	.1082
30	81	.0001	.8447	.0760	.0001	.8584	.0733
10	∞	.0000	.5622	.2775	0	.5764	.2734
20	∞	.0000	.7245	.1644	0	.7342	.1625
30	∞	.0000	.7985	.1141	0	.8070	.1141

Table 6: Exact Coverage Probabilities of 90% Confidence Intervals for Relative Risk

			Newcombe			Adjustment 1			Adjustment 2		
n_1	n_2	Max Risk	Min Cov	Mean Cov	MSE	Min Cov	Mean Cov	MSE	Min Cov	Mean Cov	MSE
10	10	99	.0001	.7655	.0673	.0091	.8768	.0089	.6648	.9156	.0019
10	20	99	.0000	.7509	.0739	.0095	.8601	.0054	.7374	.9037	.0012
20	20	99	.0167	.8302	.0344	.0332	.8928	.0029	.7881	.9100	.0008
40	40	99	.2092	.8686	.0132	.1096	.8986	.0009	.8095	.9061	.0003
50	40	99	.2847	.8770	.0098	.1307	.9021	.0007	.8609	.9081	.0002
50	50	99	.2749	.8763	.0099	.1559	.9010	.0005	.8609	.9068	.0002
75	75	99	.4217	.8869	.0044	.2796	.9006	.0003	.8721	.9043	.0002
75	100	99	.4476	.8864	.0044	.3340	.8992	.0002	.8721	.9035	.0001
50	100	99	.3056	.8744	.0101	.3892	.8959	.0003	.8579	.9043	.0001
100	100	99	.5394	.8918	.0025	.3994	.9012	.0001	.8590	.9039	.0001

Table 7: Exact Coverage Probabilities of 95% Confidence Intervals for Relative Risk

			Newcombe			Adjustment 1			Adjustment 2		
n_1	n_2	Max Risk	Min Cov	Mean Cov	MSE	Min Cov	Mean Cov	MSE	Min Cov	Mean Cov	MSE
10	10	99	.0005	.8030	.0737	.0167	.9194	.0084	.7374	.9530	.0012
10	20	99	.0000	.7859	.0808	.0956	.9022	.0062	.7374	.9411	.0010
20	20	99	.0952	.8731	.0365	.0591	.9383	.0027	.8179	.9521	.0006
40	40	99	.2445	.9142	.0153	.1096	.9469	.0007	.8822	.9528	.0001
50	40	99	.3074	.9232	.0114	.1617	.9493	.0006	.9086	.9535	.0001
50	50	99	.3062	.9221	.0116	.1560	.9477	.0005	.9086	.9522	.0001
75	75	99	.4675	.9345	.0053	.2802	.9489	.0002	.9117	.9517	.0001
75	100	99	.4670	.9338	.0054	.5268	.9474	.0001	.9192	.9508	.0000
50	100	99	.3056	.9194	.0120	.3944	.9423	.0003	.9096	.9492	.0000
100	100	99	.5588	.9397	.0031	.4015	.9495	.0001	.9192	.9515	.0000

Table 8: Exact Coverage Probabilities of 99% Confidence Intervals for Relative Risk

			Newcombe			Adjustment 1			Adjustment 2		
n_1	n_2	Max Risk	Min Cov	Mean Cov	MSE	Min Cov	Mean Cov	MSE	Min Cov	Mean Cov	MSE
10	10	99	.0146	.8329	.0774	.0169	.9585	.0072	.8895	.9845	.0004
10	20	99	.0018	.8192	.0842	.0956	.9485	.0058	.9044	.9778	.0005
20	20	99	.1331	.9065	.0382	.0607	.9733	.0021	.9265	.9866	.0002
40	40	99	.2937	.9495	.0164	.1908	.9845	.0005	.9393	.9877	.0001
50	40	99	.3689	.9602	.0114	.3908	.9872	.0003	.9106	.9863	.0001
50	50	99	.3685	.9514	.0115	.2642	.9861	.0003	.9106	.9884	.0000
75	75	99	.4950	.9717	.0059	.4429	.9875	.0001	.9603	.9889	.0000
75	100	99	.4925	.9709	.0060	.5657	.9865	.0001	.9603	.9881	.0000
50	100	99	.3950	.9570	.0117	.3960	.9722	.0003	.9582	.9861	.0001
100	100	99	.6145	.9780	.0072	.5950	.9885	.0001	.9729	.9893	.0000

Table 9: Widths of the 95% Confidence Intervals for Odds Ratio

5 Number Summaries										
	Newcombe					Adjustment				
n	Min	Q_1	Med	Q_3	Max	Min	Q_1	Med	Q_3	Max
10	.192	1.955	13.134	177.242	∞	.939	4.93	17.353	96.988	1022.5
20	.043	1.441	7.743	104.512	∞	.207	2.248	9.131	54.318	3669.4
30	.018	1.071	5.578	51.171	∞	.089	1.539	6.3073	39.559	7930.5

Table 10: Widths of the 95% Confidence Intervals for Relative Risk

5 Number Summaries							
Method	n_1	n_2	Min	Q_1	Med	Q_3	Max
Newcombe	20	20	.019	.551	1.10	3.59	∞
	40	40	.010	.393	.767	2.39	∞
	50	40	.003	.361	.727	2.35	∞
	75	100	.021	.279	.518	1.51	∞
	50	100	.001	.332	.587	1.62	∞
	100	100	.001	.249	.480	1.45	∞
Adjustment 1	20	20	.029	.582	1.12	3.54	629.8
	40	40	.011	.403	.781	2.36	1267.2
	50	40	.003	.365	.730	2.27	1264.2
	75	100	.024	.284	.525	1.51	3183.7
	50	100	.002	.341	.602	1.63	3194.1
	100	100	.003	.251	.483	1.44	3178.5
Adjustment 2	20	20	.189	.613	1.17	3.58	631.9
	40	40	.096	.412	.789	2.38	1268.3
	50	40	.088	.376	.745	2.29	1271.3
	75	100	.046	.286	.528	1.51	3173.8
	50	100	.061	.342	.601	1.63	3163.8
	100	100	.039	.254	.485	1.44	3178.9

APPENDIX B: PROGRAMS

```
%MatLab Program 1

%Exact coverage for the odds ratio

%exactcovoddsratio.m

clear

covadj = [];

covnew = [];

d = [];

for n = 30

alpha = .05;

z = icdf('norm',1-alpha/2,0,1);

n1 = n + 1;

x1 = (0:n)';

x2 = (1:n)';

x3 = (1:n)';

x4 = (1:n)';

f1 = kron(x1,ones(n^3,1));

f2 = kron(ones(n1,1),kron(x2,ones(n^2,1)));

f3 = kron(ones(n1*n,1),kron(x3,ones(n,1)));

f4 = kron(ones(n1*n^2,1),x4);

F = [f1,f2,f3,f4];

t = find(sum(F')== n);

F1 = F(t,:);

t1 = find(F1(:,1)==0);
```

```
F2 = [F1(t1,2), F1(t1,3), F1(t1,4), F1(t1,1)];

F3 = [F1(t1,3), F1(t1,4), F1(t1,1), F1(t1,2)];

F4 = [F1(t1,4), F1(t1,1), F1(t1,2), F1(t1,3)];

Fnew = [F1;F2;F3;F4];

f1 = Fnew(:,1);

f2 = Fnew(:,2);

f3 = Fnew(:,3);

f4 = Fnew(:,4);

estodds1 = (f1).*(f4)./((f2).*(f3));

estodds2 = (f1+.5).*(f4+.5)./((f2+.5).*(f3+.5));

se1 = sqrt(1./(f1)+1./(f2)+1./(f3)+1./(f4));

se2 = sqrt(1./(f1 + .5) + 1./(f2 + .5) + 1./(f3 + .5) + 1./(f4 + .5));

lb1 = estodds2.*exp(-z*se2);

ub1 = estodds2.*exp(z*se2);

se3 = 2*asinh(z*se1/2);

lb2 = estodds1.*exp(-se3);

ub2 = estodds1.*exp(se3);

w1 = find(f1 == 0);

lb2(w1) = 0;

ub2(w1) = z^2 * f4(w1)./ (f2(w1).* f3(w1));

w2 = find(f2 == 0);

lb2(w2) = f1(w2).*f4(w2)./ (f3(w2).* z^2);

ub2(w2) = inf;
```

```
w3 = find(f3 == 0);
lb2(w3) = f1(w3).*f4(w3)./ (f2(w3).* z^2);
ub2(w3) = inf;
w4 = find(f4 == 0);
lb2(w4) = 0;
ub2(w4) = z^2 * f1(w4)./ (f2(w4).* f3(w4)) ;
for i = 1:2000
    p1 = unifrnd(0,1,1,1);
    p2 = unifrnd(0,1-p1,1,1);
    p3 = unifrnd(0,1-(p1+p2),1,1);
    p4 = 1-(p1+p2+p3);
    %p1 = .1;
    %p2 = .4;
    %p3 = .4;
    %p4 = .1;
oddsratio = p1*p4/(p2*p3);
d = [d,oddsratio];
%width1 = ub1 - lb1;
%width2 = ub2 - lb2;
indodds1 = (lb1 < oddsratio & oddsratio < ub1);
indodds2 = (lb2 < oddsratio & oddsratio < ub2);
prob = gamma(n1)./(gamma(f1+1).*gamma(f2+1).*gamma(f3+1).*gamma(f4+1)).
       *(p1.^f1).*(p2.^f2).*(p3.^f3).*(p4.^f4);
```

```
totodds1 = sum(prob.*indodds1);

totodds2 = sum(prob.*indodds2);

covadj = [covcor, totodds1];

covnew = [covnew, totodds2];

end

c1 = find(1/2.25 <= d & d <= 2.25);

c1 = find(d >= 16 | d <= 1/16);

[min(covadj(c1)),mean(covadj(c1)),var(covadj(c1))+(mean(covadj(c1))-(1-alpha))^2]

[min(covnew(c1)),mean(covnew(c1)),var(covnew(c1))+(mean(covnew(c1))-(1-alpha))^2]

%plot(d,covcor(c1),'.')

end
```

```
% MatLab Program 2

% Exact coverage for Relative Risk

% Exactcovrelativerisk.m

clear

tic;

n1 = 100;

n2 = 100;

n= n1+n2;

alpha = .01;

z = icdf('norm',1-alpha/2,0,1);

covadjwoolf = [];

covnew = [];

d = [];

y1 = 0:1:n1;

y2 = 0:1:n2;

y1 = kron(y1',ones(n2+1,1));

y2 = kron(ones(n1+1,1),y2');

Y = [y1,y2];

t1 = find(y1 ~=0 | y2 ~= 0);

y3 = Y(t1,1);

y4 = Y(t1,2);

phat1 = (y3)./(n1);

phat2 = (y4)./(n2);
```

```
adjphat1 = (y1+.5)./(n1);
adjphat2 = (y2+.5)./(n2);
%corphat1 = (y1+.5)./(n1+1);
%corphat2 = (y2+.5)./(n2+1);
prelhat = (phat1)./(phat2);
adjprelhat = (adjphat1)./(adjphat2);
se1 = sqrt(1./y3 + 1./y4 - 1/n1 - 1/n2);
se2 = sqrt(1./(y1+.5) + 1./(y2+.5) - 1/(n1) - 1/(n2));
%se2 = sqrt(1./(y1+.5) + 1./(y2+.5) - 1/(n1+1) - 1/(n2+1));
se3 = 2*asinh(z*se1/2);
lbadjwoolf = adjprelhat.*exp(-z*se2);
ubadjwoolf = adjprelhat.*exp(z*se2);
lbnew = prelhat.*exp(-se3);
ubnew = prelhat.*exp(se3);
w1 = find(y3 == 0);
lbnew(w1) = 0;
ubnew(w1) = ((z^2)/(n1))/((y4(w1))./(n2));
w2 = find(y4 == 0);
lbnew(w2) = (y3(w2)./(n1))/((z^2)/(n2));
ubnew(w2) = inf;
%for i = 1:1000
    %ps = random('unif',0,1,2,1);
    %p1 = ps(1);
```

```
    %p2 = ps(2);
for p1 = .01:.01:.99
    for p2 = p1:.01:.99
 %p1 = .1;
 %p2 = .9;
        relativerisk = p1 / p2;
        d = [d,relativerisk];
%width1 = ubnew - lbnew;
%width2 = ubadjwoolf - lbadjwoolf;
indadjwoolf = (lbadjwoolf < relativerisk & relativerisk < ubadjwoolf);
indnew = (lbnew < relativerisk & relativerisk < ubnew);
prob1 = pdf('bino',y1,n1,p1).*pdf('bino',y2,n2,p2);
prob2 = pdf('bino',y3,n1,p1).*pdf('bino',y4,n2,p2);
totadjwoolf = sum(prob1.*indadjwoolf);
covadjwoolf = [covadjwoolf,totadjwoolf];
totnew = sum(prob2.*indnew);
covnew = [covnew,totnew];
end
end
[min(covadjwoolf),mean(covadjwoolf),var(covadjwoolf)+
        (mean(covadjwoolf)-(1-alpha))^2]
[min(covnew),mean(covnew),var(covnew)+(mean(covnew)-(1-alpha))^2]
```

VITA

Troy A. Bowman

Personal Data: Date of Birth: August 24, 1974
Place of Birth: Bulls Gap, Tennessee
Marital Status: Single

Education: East Tennessee State University, Johnson City, Tennessee;
Biology, Mathematics, B.S. 1998.
East Tennessee State University, Johnson City, Tennessee;
Mathematics, M.S. 2002.

Professional
Experience: Math Lab Tutor, East Tennessee State University,
January, 2000-December, 2000
Graduate Teaching Assistant, East Tennessee State University,
January, 2001-May, 2002
Math Lab Coordinator, East Tennessee State University,
June, 2002-August, 2002
Adjunct Professor, Northeast State Technical Community College,
Blountville, Tennessee, June, 2002-August, 2002