5-2011

Statistical inference of a measure for two binomial variates

Serena Petersen
University of Nevada, Las Vegas

STATISTICAL INFERENCE OF A MEASURE

FOR TWO BINOMIAL VARIATES

by

Serena Petersen

Bachelor of Science
University of Nevada, Las Vegas
2006

A thesis submitted in partial fulfillment
of the requirements for the

Master of Science in Mathematical Sciences
Department of Mathematical Sciences
College of Science

Graduate College
University of Nevada, Las Vegas
May 2011

THE GRADUATE COLLEGE

We recommend the thesis prepared under our supervision by

Serena Petersen

entitled

Statistical Inference of a Measure for Two Binomial Variates

be accepted in partial fulfillment of the requirements for the degree of

Master of Science in Mathematical Sciences

Hokwon Cho, Committee Chair

Malwane Ananda, Committee Member

Sandra Catlin, Committee Member

Chad Cross, Graduate Faculty Representative

Ronald Smith, Ph. D., Vice President for Research and Graduate Studies
and Dean of the Graduate College

May 2011

ABSTRACT

Statistical Inference of a Measure for Two Binomial Variates

by

Serena Petersen

Hokwon Cho, Ph.D., Examination Committee Chair
Associate Professor of Mathematical Sciences
University of Nevada, Las Vegas

We study measures of a comparison for two independent binomial variates which frequently occur in real situations. An estimator for measure of reduction (MOR) is considered for two sample proportions based on a modified maximum likelihood estimation. We study the desirable properties of the estimator: the asymptotic behavior of its unbiasedness and the variance of the estimator. Since the measure ρ is approximately normally distributed when sample sizes are sufficiently large, one may establish approximate confidence intervals for the true value of the estimators. For numerical study, the Monte Carlo experiment is carried out for the various scenarios of two sets of samples as well as to examine its finite sample behavior. Also, we investigate the behavior of the estimates when sample sizes get large. Two examples are provided to illustrate the use of this new measure, and extended to the hypothesis testing for further statistical inference.

ACKNOWLEDMENTS

I would like to extend my sincere gratitude and appreciation to my advisor Dr. Hokwon Cho. With his guidance, unending patience and continual encouragement, I was able to accomplish and complete this project. It has been an honor and privilege to work with him on this exciting and challenging topic.

Also, I would like to thank my Committee Members: Dr. Malwane Ananda, Dr. Sandra Catlin and Dr. Chad Cross for their contribution and advice which have greatly benefited me in achieving my goal in Applied Statistics during these past years.

In particular, I am deeply grateful to my family and friends for their support and encouragement. They provided opportunity and the means for me to complete my education.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivation of the Problem

In this thesis we are concerned with a measure of reduction for two independent binomial variates. Binomial probability phenomenon is one of the most commonly occurring distributions in our lives. Consider a problem of comparing sample proportions using a ratio from two independent binomial populations. There can be several ways of making ratios using two binomial proportions. The ratio of two binomial parameters is often called the relative risk or risk ratio, which has been studied and frequently shown in cohort studies (Katz, Baptista, Azen, and Pike, 1978), medical and pharmaceutical problems (Koopman, 1984), and epidemiological problems (Bailey, 1987), and so on. Notably Noether (1957) seems to be the first to discuss a measure of effectiveness for two binomial distributions. However, for utility of measure we devised a measure of reduction. In particular, one wishes to know how much the risk has been reduced. For a more convenient way to figure out, we consider a measure of reduction rather than making a ratio of two proportions. Then, the measure of reduction is more practical to utilize in a measurement objective and can be more useful in comparison of two binomial proportions. For example,

(i) a medical research group wants to know whether a flu shot effectively reduces the flu infection rate during the flu season;

(ii) a military test group wishes to determine whether implementing an electronic jamming technique is effective in reducing the lethality or not;

(iii) a team of operation researchers wants to figure out whether implementing a new electronic method in radar detection will reduce the detection rate or not.

The goal of this thesis is to study the properties of the measure of reduction and to construct an approximate $(1-\alpha)100\%$ confidence interval of the estimator. Or in other words, how much the new condition can lower the risk compared to the existing condition.

1.2 Assumptions and Definitions

Suppose we have two independent sequences of Bernoulli trials with nonzero probabilities. Let $X_1, X_2, \ldots, X_n$ be a sequence of Bernoulli trials with probability p_0 and $Y_1, Y_2, \ldots, Y_n$ be a sequence of Bernoulli trials with probability p_1, respectively. For instance, we are interested in measuring the "degree of reduction" in two comparable binomial variates with n $(< \infty)$ trials of each outcome/category, say p_0 represents the proportion (or probability) under no treatment made and p_1 indicates the proportion (or probability) under the new treatment/condition imposed.

Definition 1.1 (Measure of Reduction; MOR) A measure ρ is called the measure of reduction (MOR) and defined by

$$\rho = \frac{p_0 - p_1}{p_0} = 1 - \frac{p_1}{p_0}, \tag{1.1}$$

where $0 < p_0 < 1$ and $0 < p_1 < 1$.

By definition, ρ is a relative figure of merit which is based on a ratio from two independent binomial proportions, as may be seen from Equation (1.1). Suppose that p_0 indicates the true proportion of a population with a certain condition and p_1 is the true proportion of a population under the newly developed condition. Customarily, p_1 is assumed to be bigger than p_0.

Then, depending upon the values of ρ, $-\infty < \rho \le 1$, we can have the following three scenarios and their interpretations for the measure of reduction;

(i) when ρ approaches one (i.e., p_1 gets close to zero) – this means that the risk of infection or lethality is completely reduced/removed.

(ii) when ρ approaches zero (i.e., p_1 is close to p_0) - this implies that no reduction is achieved.

(iii) when ρ is negative - this indicates that a certain degree of reduction is achieved.

1.3 Probability Distributions

In this section, we study some important random variables and their probability distributions that are related to two main sampling schemes.

Definition 1.2 (Binomial Distribution) A random variable X is said to have a binomial distribution with parameters n and p if the probability mass function is given by

$$P(X = x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, \ x = 0,1,\ldots,n \\ 0, \ \text{otherwise} \end{cases}, \qquad (1.2)$$

where n is the number of total trials and p is the probability of success in each trial. We denote this by $X \sim Bin(n,p)$. It can be shown that

$$E(X) = np \text{ and } Var(X) = np(1-p).$$

The binomial probability refers to the probability that a binomial experiment consisting of n trials results in exactly x successes with probability of success p in Bernoulli trials.

Consider instead of performing a given number of trials, one conducts independent Bernoulli trials, respectively, until a desired number of successes are observed and then separated. In this setting, the total number of trials required is random. This leads us to define the following probability distribution.

Definition 1.3 (Negative Binomial Distribution) A variable X is said to have the negative binomial distribution with parameters r and p if the probability mass function is given by:

$$P(X=k) = \binom{k+r-1}{r-1}(1-p)^r \, p^k, \tag{1.3}$$

where r is the number of successes and k is the required total number of trials. We denote this by $X \sim NB(r,p)$. It can be shown that

$$E(X) = r\frac{p}{(1-p)} \text{ and } Var(X) = r\frac{p}{(1-p)^2}.$$

Where r is the number of successes, p is the probability of success and n is the number of total samples. The negative binomial distribution plays an important role by estimating the sample size when an event has a relatively small probability.

Next, we introduce a probability distribution related to a sample without replacement. This is known to be a probability for an urn model.

Definition 1.4 (Hypergeometric Distribution) Suppose we draw n objects from a population size without replacement, from an urn containing N objects in total, m of which are red. Let k be the number of red objects drawn from the urn. A random variable X is said to have the hypergeometric distribution if the probability mass function is given by:

$$P(X=k)=\frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}. \tag{1.4}$$

We denote this by $X \sim HypGeo(k;N,m,n)$. It can be shown that

$$E(X)=n\left(\frac{m}{N}\right) \text{ and } Var(X)=n\left(\frac{m}{N}\right)\left(\frac{N-m}{N}\right).$$

We note that if N is sufficiently large, then $\frac{m}{n} \approx p$ and $\frac{N-m}{N} \approx 1-p=q$ as in the binomial distribution.

1.4 Sampling Schemes

In comparing two independent populations, we use two independent samples of size n, respectively. Basically, two types of sampling schemes are feasible: fixed and sequential. In this subsection, we briefly describe the sampling schemes that are most frequently used.

We start with a method of sampling that is most commonly used in statistics. The fixed sampling method is a sampling from a population having a fixed number of n trials decided on in advance. However, problems may occur with this method because there may not have been enough measurements to obtain a desired statistical significance. This is remedied with some kind of sequential sampling.

The sequential sampling method has an advantage to optimize the sample size with respect to the objectives or inferential goals (see Cho, 2007). This sampling method is used when a sample size of n trials is not fixed in advance. The sample size is determined by the sampling results. The idea is to draw a preliminary sample of observations to determine how large the total sample size should be.

The sequential sampling technique consists of continuing the number of trials until a predetermined number, m, of distinct observations, appear in the sample with or without replacement. For example, this sampling method can be used when a true proportion is small, and a larger sample is needed in order to estimate the probability p with a specified relative error r.

Stopping rules are required for sequential sampling or inverse type sampling. A stopping rule depends on the goal of an experiment or on those conducting the experiment.

There are cases when it is difficult to obtain many samples. One way to alleviate a small sample size in an experiment is to use a resampling technique. We briefly introduce further investigation in resampling techniques and will mention this in Section 4.3 Future Research.

A resampling technique may be applied when there is a small sample size to achieve more statistical confidence. Resampling allows estimating the precision of sample statistic means, variances, medians, etc. Resampling procedures are highly computer-intensive with the most commonly used resampling techniques being jackknife and bootstrapping.

The Jackknife Method was originally designed by Quenouille (1956) and reduces the bias of an estimator. This method is used in statistical inference to estimate the bias and standard error in a statistic when a random sample of observations is used to calculate it. The jackknife estimator recalculates the statistic estimate; it leaves out one observation at a time from the sample set. From this new set of observations, an estimate for the bias and an estimate for the variance of the statistic is calculated.

The Bootstrap is a procedure for estimating the distribution of a statistic based on resampling methods. Efron and Tibshirani (1994). They describe this method as a resampling method that enables one to better understand the characteristics of an estimator without the aid of additional probability modeling. This procedure is based on resampling and simulations where we take a random sample from the sample. This new sample is taken by sampling with replacement, or in other words, some of the original sample can appear more than once. The new collection is called the bootstrap sample and is used to assess an estimator’s variability and bias, predictive performance, and significance of a test.

CHAPTER 2

FORMULATION OF THE PROBLEM

2.1 Point Estimation of ρ

Suppose we have two samples of size n ($< \infty$) from two independent populations with probabilities p_0 and p_1, respectively, say $X_1, X_2,.., X_n$ and $Y_1, Y_2,.., Y_n$. Define

$$R = \sum_{i=1}^{n} X_i \text{ and } S = \sum_{i=1}^{n} Y_i .$$

Then, $\sum_{i=1}^{n} X_i$ follows the binomial distribution with parameters of n and p_0, and $\sum_{i=1}^{n} Y_i$ follows the binomial distribution with parameters of n and p_1. That is, $R \sim \text{Bin}(n,p_0)$ and $S \sim \text{Bin}(n,p_1)$.

Definition 2.1 (Unbiasedness) An estimator $\hat{\theta}_n$, a function of R and S, is said to be unbiased for the parameter θ if

$$E\left(\hat{\theta}_n\right) = E\left[f\left(R,S\right)\right] = \theta. \tag{2.1}$$

Definition 2.2 (Bias) The bias B for an estimator $\hat{\theta}_n$, of the parameter θ, is defined by

$$B\left(\theta\right) = E\left(\hat{\theta}_n\right) - \theta. \tag{2.2}$$

Definition 2.3 (Asymptotic Unbiasedness) An estimator $\hat{\theta}_n$ is said to be asymptotically unbiased for the parameter θ if

$$E\left(\hat{\theta}_n\right) \to \theta_n \text{ as } n \to \infty. \tag{2.3}$$

From Equation (1.1), the estimator for measure of reduction for two sample proportions $\hat{p}_0$ and $\hat{p}_1$, $\hat{\rho}$, is:

$$\hat{\rho} = \frac{\hat{p}_0 - \hat{p}_1}{\hat{p}_0} = 1 - \frac{\hat{p}_1}{\hat{p}_0}, \tag{2.1}$$

where $\hat{p}_0 = \dfrac{R}{n}$ and $\hat{p}_1 = \dfrac{S}{n}$.

Since there does not exist an unbiased estimator of the measure ρ, we use the modified $\hat{\rho}$ to avoid the case of undefined $\hat{\rho}_n$ when $R = 0$:

$$\hat{\rho}_n = \frac{(R+\varepsilon)-S}{R+\varepsilon} = 1 - \frac{S}{R+\varepsilon}, \quad -\infty < \hat{\rho}_n \le 1 \tag{2.2}$$

where ε (0 < ε < 1) is an auxiliary constant. For practical purpose, one may take $\varepsilon = ½$ (e.g., see Bailey, 1987, and Cho, 2007).

2.2 Properties of the Estimator

In this section we study the fundamental properties based on first two moments of the estimator $\hat{\rho}_n$ for further investigation.

2.2.1 Expectations and Bias

Now consider the expectation of the estimator $\hat{\rho}_n$ and its bias. By definitions and independence, we have

$$\begin{aligned} E(\hat{\rho}_n) &= E\left(1 - \frac{S}{R+1/2}\right) = 1 - E\left(\frac{S}{R+1/2}\right) \\ &= 1 - E(S)E\left(\frac{1}{R+1/2}\right) = 1 - np_1 E\left(\frac{1}{R+1/2}\right). \end{aligned} \tag{2.3}$$

But, noting that (see also Cho, 2007)

$$
\begin{aligned}
E\left(\frac{1}{R+1/2}\right) &= E\left(\frac{1}{np_0+R-np_0+1/2}\right) \\
&= \frac{1}{np_0} E\left[\left(1+\frac{R-np_0+1/2}{np_0}\right)^{-1}\right] \\
&= \frac{1}{np_0} E\left[1-\left(\frac{R-np_0+1/2}{np_0}\right)+\left(\frac{R-np_0+1/2}{np_0}\right)^2+\ldots\right] \qquad (2.4)\\
&= \frac{1}{np_0} E\left[1-\frac{1}{2np_0}+\frac{np_0\left(1-p_0\right)}{\left(np_0\right)^2}+\frac{1}{4\left(np_0\right)^2}+\ldots\right] \\
&= \frac{1}{np_0}-\frac{1}{2\left(np_0\right)^2}+\frac{np_0\left(1-p_0\right)}{\left(np_0\right)^3}+\frac{1}{4\left(np_0\right)^3}+\ldots
\end{aligned}
$$

From Equations (2.3) and (2.4), we have

$$
E\left(\hat{\rho}_n\right)=1-np_1\left[\frac{1}{np_0}-\frac{1}{2\left(np_0\right)^2}+\frac{np_0\left(1-p_0\right)}{\left(np_0\right)^3}+\frac{1}{4\left(np_0\right)^3}+\ldots\right]. \qquad (2.5)
$$

Then, the bias $\hat{B}$ becomes

$$
\begin{aligned}
\hat{B} &= E(\hat{\rho}_n)-\rho \\
&= -\frac{p_1}{2np_0^2}+\frac{p_1\left(1-p_0\right)}{np_0^2}+O\left(n^{-2}\right) \qquad (2.6)\\
&= \frac{p_1\left(-p_0+1/2\right)}{np_0^2}+O\left(n^{-2}\right).
\end{aligned}
$$

Thus, $\hat{B}^2 = O(n^{-2})$ and can be neglected in the expansion.

For $n \to \infty$, it follows from Equation (2.5) that:

$$E\left(\hat{\rho}_n\right) = 1 - \frac{np_1}{np_0} \simeq \rho.$$

That is, $\hat{\rho}_n$ is asymptotically unbiased estimator of ρ.

2.2.2 Maximum Likelihood Estimation

To find the asymptotic variance of $\hat{\rho}_n$, we consider using the maximum likelihood estimates of ρ and p_0. Since $p_1 = (1-\rho)\, p_0$, the likelihood function of ρ and p_0, denoted by $L(\rho, p_0)$ is given by

$$\begin{aligned}
L(\rho, p_0) &= \binom{n}{r} p_0^r \left(1-p_0\right)^{n-r} \binom{n}{s} \left\{p_0\left(1-\rho\right)\right\}^s \left\{1-p_0\left(1-\rho\right)\right\}^{n-s} \\
&= \binom{n}{r}\binom{n}{s} p_0^{r+s} \left(1-p_0\right)^{n-r} \left(1-\rho\right)^s \left(1-p_0+\rho p_0\right)^{n-s} \qquad (2.7) \\
&\propto p_0^{r+s} \left(1-p_0\right)^{n-r} \left(1-\rho\right)^s \left(1-p_0+\rho p_0\right)^{n-s}.
\end{aligned}$$

Taking logarithm for both sides and denote the log-likelihood of $L(\rho, p_0)$ by $l(\rho, p_0)$, Equation (2.7) is

$$\begin{aligned}
l(\rho, p_0) \propto\ & (r+s)\log_e p_0 + (n-r)\log_e\left(1-p_0\right) \\
& + s\log_e \rho + (n-s)\log_e\left(1-\rho p_0\right)
\end{aligned} \qquad (2.8)$$

By setting $\frac{\partial l(\rho, p_0)}{\partial \rho} = 0$, the maximum likelihood estimate (MLE) of ρ, $\hat{\rho}_{MLE}$ is found which yields:

$$\hat{\rho}_{MLE} = 1 - \frac{s}{np_0}.$$

And letting $\frac{\partial l(\rho, p_0)}{\partial p_0} = 0$ yields an MLE of p_0, which is

$$\hat{p}_0 = \frac{r}{n}.$$

It should be noted that since

$$E(\hat{\rho}_{MLE}) = 1 - E\left(\frac{S}{np_0}\right) = 1 - \frac{1}{np_0}E(S) = 1 - \frac{1}{np_0}(np_1) = 1 - \frac{p_1}{p_0} = \rho \qquad (2.9)$$

the MLE of ρ, $\hat{\rho}_{MLE}$ is an unbiased estimator. And so is $\hat{p}_0$ because

$$E(\hat{p}_0) = E\left(\frac{R}{n}\right) = \frac{1}{n}E(R) = p_0 \qquad (2.10)$$

2.2.3 Asymptotic Variances

To find the asymptotic variance of the MLEs, we consider the Fisher's Information about ρ, p_0 and (ρ, p_0). Using Eq. (2.8) the information about ρ is

$$I(\rho) = E\left[\left(\frac{\partial l(\rho, p_0)}{\partial \rho}\right)^2\right] = -E\left(\frac{\partial^2 l(\rho, p_0)}{\partial \rho^2}\right). \qquad (2.11)$$

First we find,

$$\frac{\partial l(\rho, p_0)}{\partial \rho} = -\frac{s}{(1-\rho)} + \frac{(n-s)\, p_0}{(1 - p_0 + \rho p_0)}, \qquad (2.12)$$

then

$$\frac{\partial^2 l(\rho, p_0)}{\partial \rho^2} = -\frac{s}{(1-\rho)^2} - \frac{(n-s)\, p_0^2}{(1 - p_0 + \rho p_0)^2}. \qquad (2.13)$$

Hence,

$$
\begin{aligned}
I(\rho) &= -E\left[-\frac{s}{(1-\rho)^2}-\frac{(n-s)\,p_0^2}{(1-p_0+\rho p_0)^2}\right] \\
&= \frac{np_1}{(1-\rho)^2}+\frac{(n-np_1)\,p_0^2}{(1-p_0+\rho p_0)^2} \\
&= \frac{np_0}{(1-\rho)}+\frac{np_0^2}{(1-p_0+\rho p_0)} \\
&= \frac{np_0}{(1-\rho)(1-p_0+\rho p_0)}.
\end{aligned}
\tag{2.14}
$$

Similarly, the information about p_0 is given by

$$
I(p_0) = E\left[\left(\frac{\partial l(\rho, p_0)}{\partial p_0}\right)^2\right] = -E\left(\frac{\partial^2 l(\rho, p_0)}{\partial {p_0}^2}\right). \tag{2.15}
$$

First we find,

$$
\frac{\partial l(\rho, p_0)}{\partial p_0} = \frac{r+s}{p_0}-\frac{(n-r)}{(1-p_0)}-\frac{(n-s)(1-\rho)}{(1-p_0+\rho p_0)}, \tag{2.16}
$$

then

$$
\frac{\partial^2 l(\rho, p_0)}{\partial {p_0}^2} = \frac{-(r+s)}{p_0^2}-\frac{(n-r)}{(1-p_0)^2}-\frac{(n-s)(1-\rho)^2}{(1-p_0+\rho p_0)^2}. \tag{2.17}
$$

Hence,

$$
\begin{aligned}
I(\rho) &= -E\left[\frac{-(r+s)}{p_0^2} - \frac{(n-r)}{(1-p_0)^2} - \frac{(n-s)(1-\rho)^2}{(1-p_0+\rho p_0)^2}\right] \\
&= \frac{n(p_0+p_1)}{p_0^2} + \frac{n(1-p_0)}{(1-p_0)^2} + \frac{n(1-p_1)(1-\rho)^2}{(1-p_0+\rho p_0)^2} \\
&= \frac{n}{p_0}\left(1+(1-\rho)+\frac{p_0}{(1-p_0)}+\frac{p_0(1-\rho)^2}{p_0(1-p_0(1-\rho))}\right) \qquad (2.18) \\
&= \frac{n}{p_0}\left(\frac{2-2p_0+2\rho p_0-\rho}{(1-p_0)(1-p_0(1-\rho))}\right) \\
&= \frac{n}{p_0}\left(\frac{1}{1-p_0}+\frac{1-\rho}{1-p_0(1-\rho)}\right).
\end{aligned}
$$

The joint information about ρ, and p_0, denoted by $I(\rho, p_0)$, is

$$I(\rho, p_0) = -E\left(\frac{\partial^2 l(\rho, p_0)}{\partial\rho\partial p_0}\right). \qquad (2.19)$$

First we find,

$$\frac{\partial l(\rho, p_0)}{\partial p_0} = \frac{r+s}{p_0} - \frac{(n-r)}{(1-p_0)} - \frac{(n-s)(1-\rho)}{(1-p_0+\rho p_0)}, \qquad (2.20)$$

then

$$\frac{\partial^2 l(\rho, p_0)}{\partial\rho\partial p_0} = -\frac{(n-s)}{(1-p_0+\rho p_0)^2}. \qquad (2.21)$$

Hence,

$$\begin{aligned} I(\rho, p_0) &= -E\left(-\frac{(n-s)}{(1-p_0+\rho p_0)^2}\right) \\ &= \frac{n-np_1}{(1-p_0(1-\rho))^2} \\ &= \frac{n(1-p_0(1-\rho))}{(1-p_0(1-\rho))^2} \\ &= \frac{n}{1-p_0(1-\rho)}. \end{aligned} \tag{2.22}$$

Then combining Eqs. (2.14), (2.18) and (2.22), we finally obtained the information matrix about (ρ,p_0),

$$\mathbf{I}(\rho, p_0) = \begin{bmatrix} E\left(\frac{\partial l(\rho, p_0)}{\partial \rho}\right)^2 & -E\left(\frac{\partial^2 l(\rho, p_0)}{\partial \rho \partial p_0}\right) \\ -E\left(\frac{\partial^2 l(\rho, p_0)}{\partial \rho \partial p_0}\right) & E\left(\frac{\partial l(\rho, p_0)}{\partial p_0}\right)^2 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{np_0}{(1-\rho)(1-p_0(1-\rho))} & \frac{n}{1-p_0(1-\rho)} \\ \frac{n}{1-p_0(1-\rho)} & \frac{n}{p_0}\left(\frac{1}{1-p_0}+\frac{1-\rho}{1-p_0(1-\rho)}\right) \end{bmatrix} \tag{2.23}$$

So,

$\mathbf{I}^{-1}(\rho, p_0) =$

$$\frac{\rho(1-p_0)(1-p_0(1-\rho))}{n}\begin{bmatrix} \frac{1}{p_0}\left(\frac{1}{1-p_0}+\frac{1-\rho}{1-p_0(1-\rho)}\right) & -\frac{1}{1-p_0(1-\rho)} \\ -\frac{1}{1-p_0(1-\rho)} & \frac{p_0}{p_1(1-p_0(1-\rho))} \end{bmatrix}. \tag{2.24}$$

Thus, from Equation (2.24) the asymptotic variance of $\hat{\rho}_{MLE}$ is

$$\begin{aligned} &Var(\hat{\rho}_{MLE}) = \\ &\frac{\rho(1-p_0)(1-p_0(1-\rho))}{n}\left[\frac{1}{p_0}\left(\frac{1}{(1-p_0)}+\frac{1-\rho}{(1-p_0(1-\rho))}\right)\right] \\ &= \frac{\rho(1+\rho-2\rho p_0)}{np_0}. \end{aligned} \tag{2.25}$$

Using Slutsky's theorem and asymptotic normality of $\hat{\rho}_{MLE}$, it follows

$$\sqrt{n}\left(\hat{\rho}_{MLE}-\rho\right) \overset{d}{\simeq} N\left(0,\sigma^2\right), \tag{2.26}$$

where $\sigma^2 = \rho(1+\rho-2\rho p_0)/p_0$.

Now we consider the asymptotic variance of the estimator, $\hat{\rho}_n = 1-\frac{S}{R+1/2}$.

$$\begin{aligned} Var\left(\hat{\rho}_n\right) &= Var\left(1-\frac{S}{R+1/2}\right) = Var\left(\frac{S}{R+1/2}\right) \\ &= E\left[\left(\frac{S}{R+1/2}\right)^2\right]-\left[E\left(\frac{S}{R+1/2}\right)\right]^2 \\ &= E\left(S^2\right)\bullet E\left(\frac{S}{R+1/2}\right)^2-\left[E(S)\bullet E\left(\frac{S}{R+1/2}\right)\right]^2. \end{aligned} \tag{2.27}$$

However, applying the same manner shown in Equation (2.4), we have (see also Cho, 2007)

$$\begin{aligned} E\left(\frac{1}{R+1/2}\right)^2 &= E\left[\frac{1}{np_0}\left(1+\frac{R-np_0+1/2}{np_0}\right)^{-1}\right]^2 \\ &= \frac{1}{(np_0)^2}E\left[1-2\left(\frac{R-np_0+1/2}{np_0}\right)+3\left(\frac{R-np_0+1/2}{np_0}\right)^2+\dots\right] \\ &= \frac{1}{(np_0)^2}\left[1-\frac{1}{np_0}+\frac{3np_0(1-p_0)}{(np_0)^2}+\frac{3}{4(np_0)^2}+\dots\right]. \end{aligned} \tag{2.28}$$

Using Equation (2.4), (2.27) and (2.28),

$$
\begin{aligned}
Var\left(\hat{\rho}_n\right) &= \left[np_1\left(1-p_1\right)+\left(np_1\right)^2\right]\left[\frac{1}{\left(np_0\right)^2}\left[1-\frac{1}{np_0}+\frac{3np_0\left(1-p_0\right)}{\left(np_0\right)^2}+\frac{3}{4\left(np_0\right)^2}+\ldots\right]\right] \\
&\quad -\left[np_1\left[\frac{1}{np_0}-\frac{1}{2\left(np_0\right)^2}+\frac{\left(1-p_0\right)}{\left(np_0\right)^2}+\frac{1}{4\left(np_0\right)^3}+\ldots\right]\right]^2 \\
&= \left[\frac{p_1\left(1-p_1\right)}{\left(np_0^2\right)}+\rho^2\right]\left[1+\frac{2-3p_0}{\left(np_0\right)}+\frac{3}{4\left(np_0\right)^2}+\ldots\right] \\
&\quad -\left(np_1\right)^2\left[\frac{1}{\left(np_0\right)^2}-\frac{1}{\left(np_0\right)^3}+\frac{2\left(1-p_0\right)}{\left(np_0\right)^3}+\frac{1}{4\left(np_0\right)^4}+\frac{\left(1-p_0\right)^2}{\left(np_0\right)^4}+\frac{\left(1-p_0\right)}{2\left(np_0\right)^4}+\ldots\right].
\end{aligned}
\tag{2.29}
$$

Simplifying,

$$
\begin{aligned}
Var\left(\hat{\rho}_n\right) &= \left[\frac{\rho\left(1-p_1\right)}{\left(np_0\right)}+\rho^2\right]\left[1+\frac{2-3p_0}{\left(np_0\right)}+\frac{3}{4\left(np_0\right)^2}+\ldots\right] \\
&\quad -\left(np_1\right)^2\left[\frac{1}{\left(np_0\right)^2}\left[1+\frac{\left(1-2p_0\right)}{\left(np_0\right)}+\frac{1}{4\left(np_0\right)^2}+\frac{\left(1-p_0\right)^2}{\left(np_0\right)^2}+\ldots\right]\right] \\
&= \left[\frac{\rho\left(1-p_1\right)}{\left(np_0\right)}+\rho^2\right]\left[1+\frac{2-3p_0}{\left(np_0\right)}+O\left(-n^2\right)\right]-\rho^2\left[1+\frac{\left(1-2p_0\right)}{\left(np_0\right)}+O\left(-n^2\right)\right] \\
&\approx \frac{\rho\left(1-p_1\right)}{\left(np_0\right)}\left(1+\frac{2-3p_0}{\left(np_0\right)}\right)+\rho^2\left(1+\frac{2-3p_0}{\left(np_0\right)}\right)-\rho^2\left(1+\frac{1-2p_0}{\left(np_0\right)}\right) \\
&= \frac{\rho\left(1-p_1\right)}{\left(np_0\right)}\left(1+\frac{2-3p_0}{\left(np_0\right)}\right)+\rho^2\left(\frac{1-p_0}{\left(np_0\right)}\right) \\
&= \frac{\rho\left(1-p_1\right)}{\left(np_0\right)}+\frac{\rho^2\left(1-p_0\right)}{\left(np_0\right)}+O\left(-n^2\right)\approx\frac{\rho\left(1+\rho-2\rho p_0\right)}{np_0}.
\end{aligned}
\tag{2.30}
$$

From Equation (2.25) to (2.30) we see the two variances of $\hat{\rho}_n$ have asymptotically the same variance. Therefore, we conclude that the two estimators of $\hat{\rho}_{MLE} = 1 - \frac{S}{R}$ and $\hat{\rho}_n = 1 - \frac{S}{R + 1/2}$, are asymptotically equivalent for sufficiently large *n*.

CHAPTER 3

NUMERICAL STUDIES AND EXAMPLES

3.1 Simulation Studies

Monte Carlo experimentation is carried out to investigate the behavior of the measure of reduction we have devised. Selected values for p_0 and p_1 were chosen to generate the data sets consisting of sequences of binomial variables based on a predetermined fixed number of trials for each case. Two sample proportions of p_0 and p_1 are computed and the point estimator of the measure of reduction ρ is also calculated with independent replications (denoted by m) of 100, 1000, 5000 and 10,000. For the interval estimation of ρ, the standard error (S.E) is also calculated along with an 95% empirical confidence interval and a 95% confidence interval around $\hat{\rho}$. The results of the Monte Carlo simulation are summarized in the following tables, which show the number of replicates, estimates of p_0, p_1 and ρ, and the lower and the upper bounds for the confidence intervals for each nominal level 95% and the 95% empirical, respectively.

In addition, we plot the empirical probability distribution of ρ, the measure of reduction, for each case to illustrate the asymptotic behavior of the measure, namely, the normal approximation to the binomial distributions in the probability distribution for the measure of reduction (MOR).

3.2 Fatality in Infectious Disease

We study an example of the vaccine effectiveness of fatality in infectious disease. Suppose that a group of researchers want to study the effectiveness of vaccine for a certain type of infectious disease such as influenza. Let p_0 be the proportion of being fatal with no vaccine treated for a population under no vaccine treated, and p_1 be the proportion of being fatal with vaccine treated.

Definition 3.1 (Reduction in Fatality; RIF) A measure of effectiveness in reduction for the infectious disease, ρ is called the reduction in fatality (RIF) and defined by

$$\rho_{RIF} = \frac{p_0 - p_1}{p_0} = 1 - \frac{p_1}{p_0}, \qquad (3.1)$$

where $0 < p_0 < 1$ and $0 < p_1 < 1$.

Suppose that the success rate of the existing vaccine is known to be 40%, i.e., $p_0 = 0.4$. A research group developed a new vaccine for the disease, which may reduce 50% (since $\rho = -0.5$), compared to the existing method to prevent the disease. After the Monte Carlo experiment set up, we summarize the results in Table 3.1. The numbers of replicates are shown in the first column, with estimates for p_0, p_1 and ρ, the standard error for each of the experiment, and corresponding approximate 95% confidence interval and an 95% empirical confidence interval. In addition, similar results are provided in 3 other scenarios, Tables 3.2-3.4.

Table 3.1: $\rho = -0.5$ with $p_0 = 0.4$ and $p_1 = 0.6$

Number of Replicates X_i's and Y_i's	$\hat{p}_0$	$\hat{p}_1$	**Estimate of ρ, $\hat{\rho}$**	**Standard Error**	**95% Confidence Interval**	**95% Empirical Confidence Interval**
100	0.4300	0.6600	−0.5618	0.0817	(−0.7219, −0.4017)	(−0.5526, −0.4091)
1000	0.4230	0.5960	−0.5428	0.0655	(−0.6712, −0.4144)	(−0.5238, −0.4872)
5000	0.4007	0.6014	−0.5377	0.0596	(−0.6545, −0.4209)	(−0.5122, −0.4878)
10000	0.4001	0.6006	−0.5242	7.87E-03	(−0.5088, −0.5396)	(−0.5106, −0.5000)

For example, based on the 10000 times from the last column in the table, we are statistically sure that the true value (percentage of reduction) lies approximately between −0.51 and −0.54 with 95% confidence.

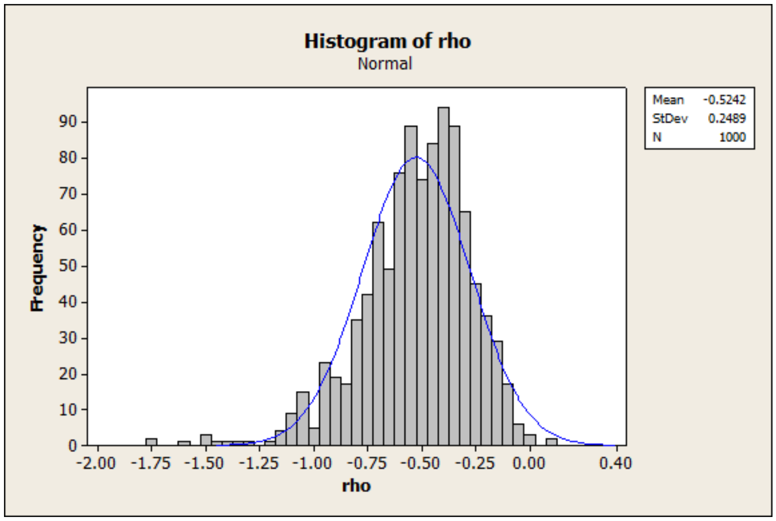

Figure 3.1a: $\rho = -0.5$ with $p_0 = 0.4$ and $p_1 = 0.6$ for 10000 Replicates

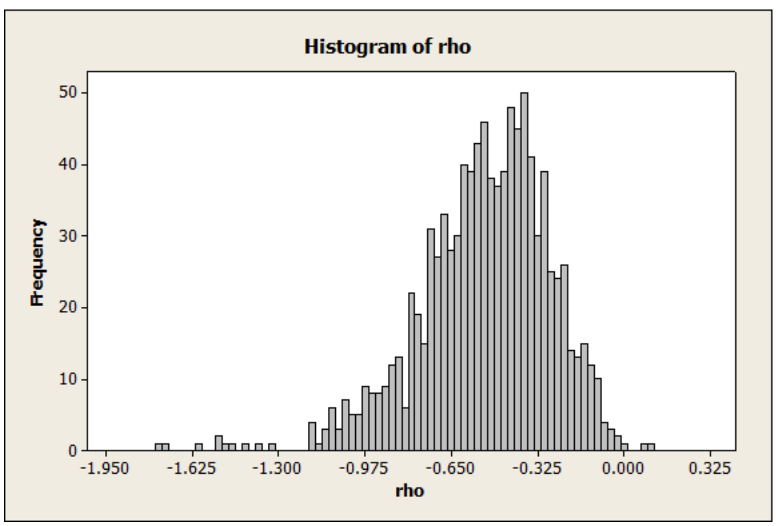

Figure 3.1b: $\rho = -0.5$ with $p_0 = 0.4$ and $p_1 = 0.6$ for 10000 Replicates

Table 3.2: $\rho = -1.0$ with $p_0 = 0.3$ and $p_1 = 0.6$

Number of Replicates X_i's and Y_i's	$\hat{p}_0$	$\hat{p}_1$	Estimate of ρ, $\hat{\rho}$	Standard Error	95% Confidence Interval	95% Empirical Confidence Interval
100	0.2589	0.5678	−1.1322	0.1470	(−1.4203, −0.8441)	(−1.0645, −0.9032)
1000	0.3330	0.5890	−1.0640	0.1070	(−1.2737, −0.8543)	(−1.0313, −0.9677)
5000	0.3034	0.6012	−1.0581	0.0752	(−1.2055, −0.9107)	(−1.0303, −0.9722)
10000	0.2995	0.5976	−1.0505	0.0117	(−1.0734, −1.0275)	(−1.0263, −0.9927)

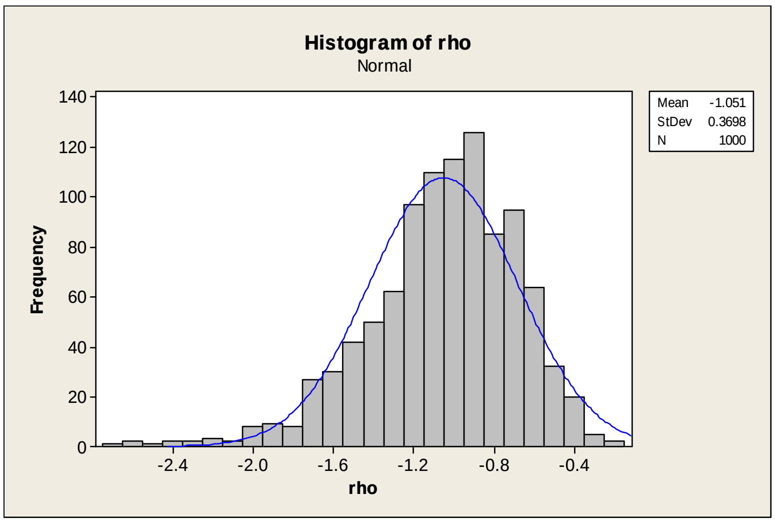

Figure 3.2a: $\rho = -1.0$ with $p_0 = 0.3$ and $p_1 = 0.6$ for 10000 Replicates

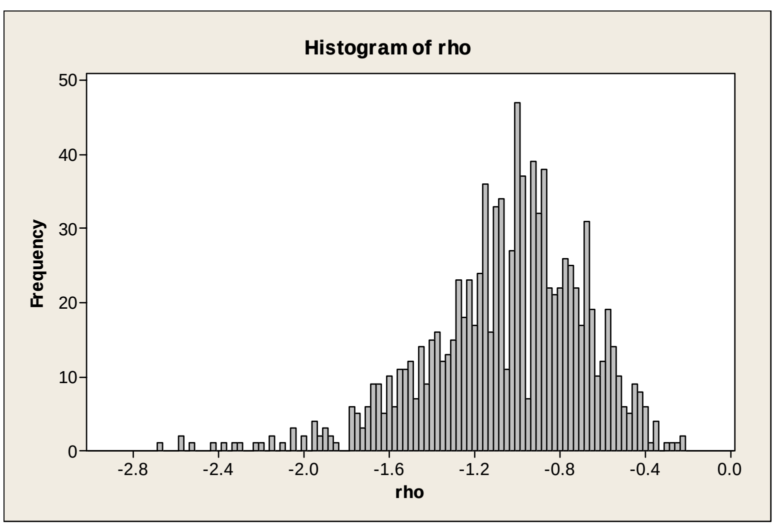

Figure 3.2b: $\rho = -1.0$ with $p_0 = 0.3$ and $p_1 = 0.6$ for 10000 Replicates

Table 3.3: $\rho = -3.0$ for $p_0 = 0.2$ and $p_1 = 0.8$

Number of Replicates X_i's and Y_i's	$\hat{p}_0$	$\hat{p}_1$	**Estimate of ρ, $\hat{\rho}$**	**Standard Error**	**95% Confidence Interval**	**95% Empirical Confidence Interval**
100	0.2500	0.7600	−3.7332	0.2530	(−4.2291, −3.2373)	(−3.1500, −2.8421)
1000	0.1830	0.7870	−3.5512	0.1470	(−3.8393, −3.2631)	(−3.0526, −2.9474)
5000	0.1932	0.7962	−3.3629	0.0748	(−3.5095, −3.2163)	(−3.0476, −2.9546)
10000	0.2003	0.8013	−3.1345	0.0292	(−3.1917, −3.0773)	(−3.0455, −2.9565)

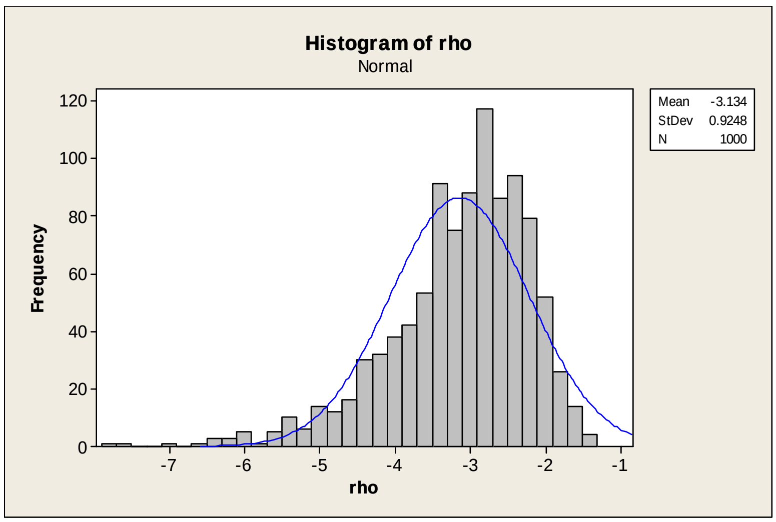

Figure 3.3a: $\rho = -3.0$ with $p_0 = 0.2$ and $p_1 = 0.8$ for 10000 Replicates

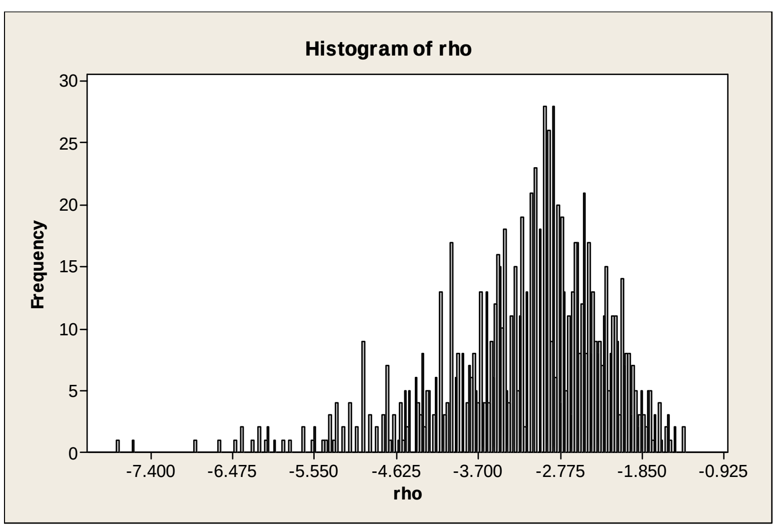

Figure 3.3b: $\rho = -3.0$ with $p_0 = 0.2$ and $p_1 = 0.8$ for 10000 Replicates

Table 3.4: $\rho = -1.0$ for $p_0 = 0.05$ and $p_1 = 0.1$

Number of Replicates X_i's and Y_i's	$\hat{p}_0$	$\hat{p}_1$	**Estimate of ρ, $\hat{\rho}$**	**Standard Error**	**95% Confidence Interval**	**95% Empirical Confidence Interval**
100	0.0425	0.0976	−1.4833	0.201	(−1.8773, −1.0893)	(−1.1667, −0.8333)
1000	0.0556	0.0984	−1.2941	0.154	(−1.5959, −0.9923)	(−1.1429, −0.8875)
5000	0.0471	0.0911	−1.1628	0.0587	(−1.2779, −1.0476)	(−1.1111, −0.9875)
10000	0.04967	0.0910	−1.0884	0.0406	(−1.1680, −1.0088)	(−1.0857, −0.9889)

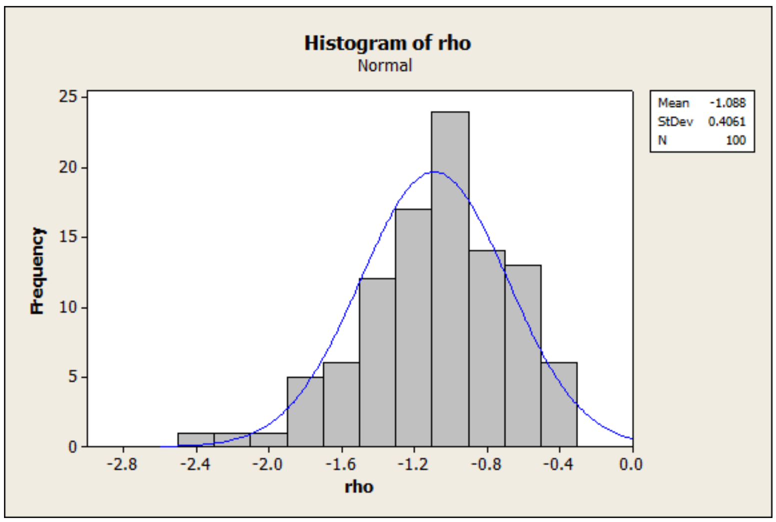

Figure 3.4a: $\rho = -1.0$ with $p_0 = 0.05$ and $p_1 = 0.1$ for 10000 Replicates

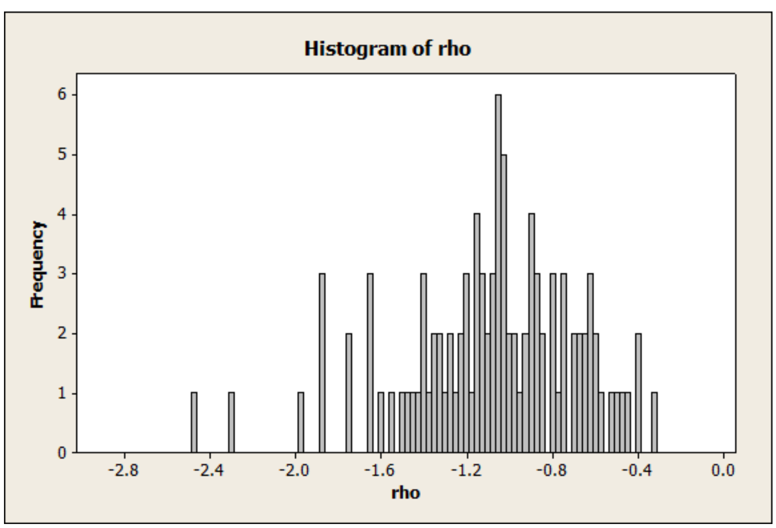

Figure 3.4b: $\rho = -1.0$ with $p_0 = 0.05$ and $p_1 = 0.1$ for 10000 Replicates

When the probability of an event p_0 (or p_1) is small, it may be desirable to use Negative Binomial sampling.

From Table 3.1 to Table 3.4 we observe that all of the Monte Carlo estimates, $\hat{p}_0$, $\hat{p}_1$, and $\hat{\rho}$, converge to the corresponding true values of parameters as n gets large. Also, the standard error (S.E.) decreases dramatically as the number of replicates increase. We surmise that the results of the Monte Carlo experiment provide the substantial amount of numerical evidence and strong belief in verifying the analytical results which are shown in the previous chapter. The estimates of p_0 and p_1 get closer to their predetermined values as the number of replicates increase. Similar results from the Monte Carlo experiment are seen with the point estimator ρ.

3.3 ECM Effectiveness

In this section, we consider the measure of reduction in the effectiveness of airborne electronic countermeasures (ECM). We study the measure of effectiveness by comparing the number of hits under two conditions whether the ECM is turned on or off. Then, the reduction in lethality is related to the proportions of being hit by a surface-to-air missile (SAM). In particular for the ECM, the measure of reduction (MOR) in effectiveness for airborne electronic countermeasures is called the reduction in lethality (RIL). The following definitions are useful when discussing ECM and SAM:

(i) a hit is the event of a SAM engaging and hitting a target,

(ii) a miss is the event where the SAM misses the engaged target,

(iii) a wet condition is when ECM is turned on, and

(iv) a dry condition is when ECM is turned off.

Definition 3.2 (Reduction in Lethality; RIL) A measure of effectiveness in reduction for the ECM, ρ is called the reduction in Lethality (RIL) and defined by

$$\rho_{RIL} = \frac{p_d - p_w}{p_d} = 1 - \frac{p_w}{p_d}, \tag{3.2}$$

where $0 < p_d < 1$ and $0 < p_w < 1$ with

p_d = proportion of being hit by missile under dry condition,

p_w = proportion of being hit by missile under wet condition.

By definition in Equation (3.2), the measure of reduction in lethality (RIL) ρ_{RIL} is a relative figure of merit rather than an absolute measurement. The advantage of the use of RIL is merely comparing the wet performance to the corresponding dry performance. For instance, if the wet proportion of hit were zero, that is $p_w = 0$, then $\rho_{RIL} = 1$. This implies that the ECM device completely removed the lethality of the missile. Similarly, if there was no difference in performance under two conditions, i.e., $p_d = p_w$, then ρ_{RIL} becomes zero. This indicates that the ECM device did not reduce the lethality of the incoming missile at all. It should be noted that the values of ρ can be negative. In fact, the target could be more vulnerable to the missile attack when the ECM device is on than when it is off in real situation.

3.3.1 ECM Measure Analysis

Requirements for ECM, to determine the effectiveness of an ECM device, lack sufficient information. Requirements often come in the form of a single observed $\hat{\rho}$

value, which a test director must use a pass or fail criteria for the ECM system in question. Also, due to the high expense of missile testing, we consider n to be small, say $n \leq 10$, noting that n is the number of runs. The methodology presented, Equation (2.1), optimizes the ECM test design for minimum cost. This will be discussed further in the next section. Note that $p_1 = p_w$ is the probability of a wet hit, $p_0 = p_d$ is the probability of a dry hit and we assume p_w and p_d to be the same throughout the experiment. A simulated example is given to illustrate what the new measure ρ would result in. The simulated example shows the measure of ρ given with independent replications of 100, 1000, 5000 and 10000, along with the expectation of $p_{0,}$ $p_{1,}$ $\hat{\rho}$, with a 95% confidence interval and 95% empirical confidence interval around $\hat{\rho}$.

Table 3.5: $\rho = -0.43$ for $p_0 = 0.56$ and $p_1 = 0.8$

Number of Replicates X_i's and Y_i's	$\hat{p}_0$	$\hat{p}_1$	Estimate of ρ, $\hat{\rho}$	Standard Error	95% Confidence Interval	95% Empirical Confidence Interval
100	0.5890	0.8213	−0.5484	0.2080	(−0.9561, −0.1407)	(−0.5179, −0.4231)
1000	0.5764	0.8184	−0.5391	0.0579	(−0.6526, −0.4256)	(−0.5098, −0.4423)
5000	0.5574	0.8056	−0.4877	0.0353	(−0.5569, −0.4185)	(−0.5087, −0.4921)
10000	0.5585	0.8018	−0.4468	0.00469	(−0.4560, −0.4376)	(−0.5047, −0.4915)

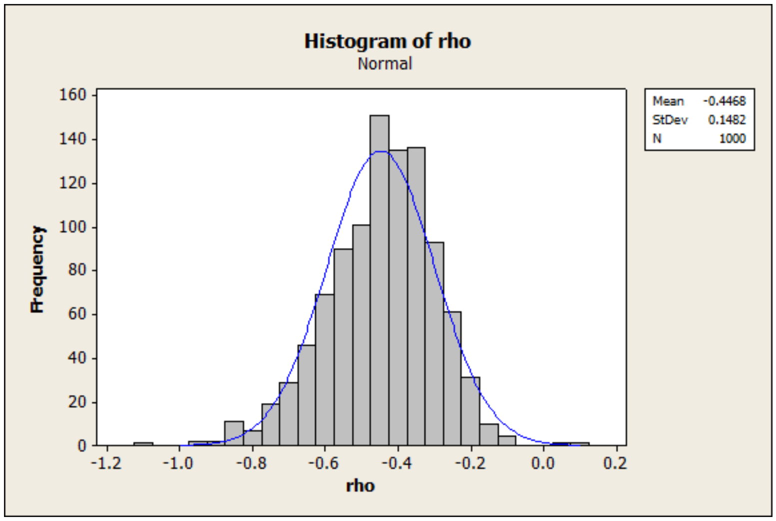

Figure 3.5a: $\rho = -0.43$ with $p_0 = 0.56$ and $p_1 = 0.8$ for 10000 Replicates

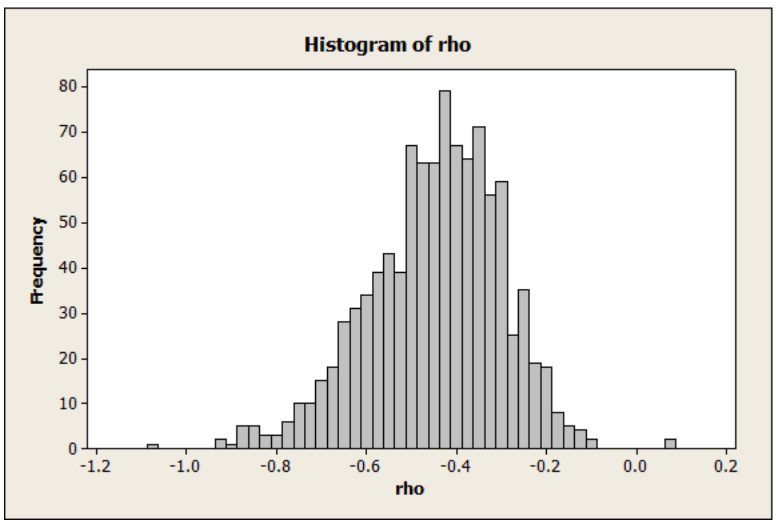

Figure 3.5b $\rho = -0.43$ with $p_0 = 0.56$ and $p_1 = 0.8$ for 10000 Replicates

In Table 3.5, we infer that the estimated proportions, $\hat{p}_0$ and $\hat{p}_1$, and the estimated measure $\hat{\rho}$ are convergent to the parameters p_0, p_1 and ρ, respectively as the number of replicates increases. It is also noticeable that the standard errors (S.E.) monotonically decrease as the sample size n gets bigger. Therefore, the above numerical evidence indicates that the small sample behavior lends support to the asymptotic behavior of the measure when the sample size gets bigger. In addition, we have verified that as n increases the shape of the plot is getting close to the shape of the normal distribution.

3.3.2 Small Sample Study in ECM

In real ECM experiment, the small sample test is frequently performed due mainly to cost reasons. In this section, we study enumerate the sample space of the joint probability mass function of $p(x_i,y_i)$. Suppose we have $n = 5$ for both shots under the wet condition and shots under the dry condition. Then the values of the measure ρ are listed in Table 3.6. When p_0 is 0, ρ is undefined. To compensate, we take $\varepsilon = 0.01$ for convenience:

$$\rho = 1 - \frac{p_1}{\left(p_0 + 0.01\right)}.$$

This ensures ρ is not undefined.

Table 3.6: The value of ρ when $n = 5$ for $p_0 = 0.56$ and $p_1 = 0.8$

$p(x,y)$		p_1: Number of hits under wet conditions					
		0	1	2	3	4	5
p_0: Number of hits under dry conditions	0	1 (undefined)	−19 (−∞)	−39 (−∞)	−59 (−∞)	−79 (−∞)	−99 (−∞)
	1	1	0	−1	−2	−3	−4
	2	1	0.500	0	−0.500	−1	−1.500
	3	1	0.667	0.333	0	−0.333	−0.667
	4	1	0.750	0.500	0.250	0	−0.250
	5	1	0.800	0.600	0.400	0.200	0

By independence, the joint pmf:

$$
\begin{aligned}
p(x_i, y_i) &= p(x_i)\, p(y_i) \\
&= \binom{5}{x} p_0^x (1-p_0)^{5-x} \binom{5}{y} p_1^y (1-p_1)^{5-y} \\
&= \left[(1-p_0)(1-p_1)\right]^5 \binom{5}{x}\binom{5}{y} p_0^x p_1^y \left(\frac{p_0}{1-p_0}\right)^x \left(\frac{p_1}{1-p_1}\right)^y .
\end{aligned}
$$

Note that the joint probability values for ρ are calculated as follows:

$$\sum_{i=0}^{5}\sum_{j=0}^{5} p(x_i, y_j) = 1.$$

The n, joint probability $p_{x,y}(x_i,y_i)$ are:

Table 3.7: ECM Joint Probabilities of $\hat{\rho}$ when $n = 5$ for $p_0 = 0.56$ and $p_1 = 0.8$

$p(x_i,y_j)$		p_1: Number of hits under wet conditions					
		0	1	2	3	4	5
p_0: Number of hits under dry conditions	0	5.277E-06	0.0001	0.0008	0.0034	0.0068	0.0054
	1	3.358E-05	0.0007	0.0054	0.0215	0.0430	0.0344
	2	8.548E-05	0.0017	0.0137	0.0547	0.1094	0.0875
	3	1.088E-04	0.0022	0.0174	0.0696	0.1393	0.1114
	4	6.923E-05	0.0014	0.0111	0.0443	0.0886	0.0709
	5	1.762E-05	0.0004	0.0028	0.0113	0.0226	0.0180

Then, we plot the probabilities of the measure of reduction, RIL, ρ.

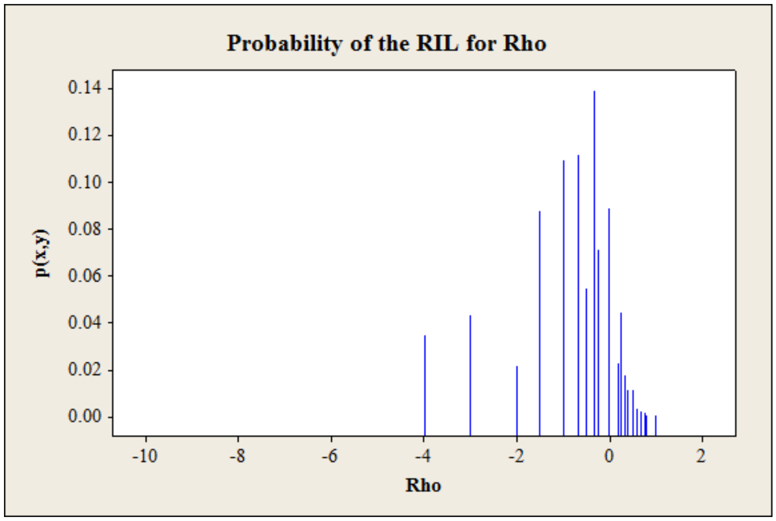

Figure 3.6: Plot for probability of ρ when n = 5 for p_0 = 0.56 and p_1 = 0.8

Furthermore, we know that $\hat{\rho}$ is approximately normally distributed. For large n, there exists a UMP test (See Mood, Graybill and Boes, 1974). Suppose one may wish to test the RIL, ρ, H_0: $\rho \leq \rho_0$ versus H_1: $\rho > \rho_0$, where ρ_0 is the desirable threshold that has been predetermined. We know that $\hat{\rho} \sim N(\rho,\ \sigma_\rho^2)$ noting that

$$\frac{\hat{\rho} - \rho_0}{\sigma_\rho / \sqrt{n}} \sim Z.$$

CHAPTER 4

CONCLUSION

4.1 Concluding Remarks

We have studied a relative measure of reduction based on a ratio of the two binomial variates. We also investigated the desirable properties of the measure, such as unbiasedness and efficiency. The numerical analysis, through Monte Carlo experiments, show that the measure ρ we devised is useful and easy to understand. Also, the numerical study illustrates how effectively used the measure is for both the infectious disease and for ECM.

4.2 Future Research

In ECM cases, there is a high cost and a small amount of opportunity to run each experiment. Again, one way to mitigate this difficulty is to apply a resampling technique. The resampling allows one to estimate the precision of sample statistic means, variances, medians, etc. If we use different types of resampling schemes we are able to optimize the sample sizes from the given lower limits and upper limits of the confidence intervals. In other words, an experiment may achieve more statistical confidence when these techniques are applied.

REFERENCES

[1] Bailey, B. (1987). Confidence Limits to the Risk Ratio. *Biometrics*, **43**, 201-205.

[2] Casella, G. and Berger, R. (2002) *Statistical Inference*. 2nd Edition. Duxbury, Pacific Grove.

[3] Cho, H. (2007). Sequential risk-efficient estimation for the ration of two binomial proportions. *Journal of Statistical Planning and Inference*, **137**, 2336-2346.

[4] Efron, B and Tibshirani, R. (1994). *An Introduction to the Bootstrap*. Chapman & Hall. CRC.

[5] Evans, M., Tinsley, R. and Cho, H. (2002) Sample size selection for ECM effectiveness testing. Unpublished manuscript.

[6] Fleiss, Joseph L. (1981) *Statistical Methods for Rates and Probabilities*. 2nd Edition. Wiley, NewYork.

[7] Govindarajulu, Z. (1999) *Elements of Sampling Theory and Methods*. Prentice Hall, New Jersey.

[8] Hogg, R., and Craig, A. (1994). *Introduction to Mathmatecal Statistics*. 5th Edition. Upper Saddle River, New Jersey.

[9] Koopman, P. (1984). Confidence intervals for the ratio of two binomial proportions. *Biometrics*, **40**, 513-517.

[10] Katz, D., Baptista, J., Azen, S., and Pike, M. (1978). Obtaining confidence intervals for the risk ratio in cohort studies. *Biometrics* **34**, 469-474.

[11] Kvam, P. and Vidakovic, B. (2007) *Nonparametric Statistics with Applications to Science and Engineering*. Wiley, New Jersey.

[12] Mood, A., Graybill, F. and Boes, D. (1974) *Introduction to the Theory of Statistics*. 3rd Edition. Boston, MA.

[13] Mooney, C. and Duval, R. (1993) *Bootstrapping a Nonparametric Approach to Statistical Inference*. SAGE University Paper, Newbury Park.

[14] Noether, G. (1957) Two confidence intervals for the ratio of two probabilities and some measures of effectiveness. *American Statistical Association Journal*, **52**, 36-45.

[15] Quenouille, M. (1956). Notes on bias in estimation. *Biometrika*, **61**, 353-360.

[16] Stimson, G. W. (1998) *Introduction to Airborne Radar*. 2^{nd} Edition. SciTech Publishing, North Carolina.

VITA

Graduate College
University of Nevada, Las Vegas

Serena Petersen

Degrees:
Bachelor of Science, Mathematical Sciences, 2006
University of Nevada, Las Vegas

Thesis Title: Statistical Inference of a Measure for Two Binomial Variates

Thesis Examination Committee:
Chairperson, Hokwon Cho, Ph. D.
Committee Member, Malwane Ananda, Ph. D.
Committee Member, Sandra Catlin, Ph. D.
Graduate Faculty Representative, Chad Cross, Ph. D.